



# Multicast

---



Redes de Comunicação de Dados

---



- Introdução
- Grupos Multicast: IGMP
- Encaminhamento Multicast
  - Distance Vector Multicast Routing Protocol
  - Core Base Trees
  - Protocol Independent Multicast
  - Border Gateway Management Protocol

# Introdução ao modelo IP Multicast

---



- Distribuição eficiente de dados no modelo de “um para muitos”
    - Distribuição em forma de árvores
    - Os pacotes atravessam as ligações apenas uma vez
  - Endereçamento é independente da localização
    - Endereço IP por grupo multicast
  - Modelo de serviços orientado ao terminal
    - As aplicações podem subscrever ou abandonar grupos *multicast*
    - Os emissores não sabem quem recebem os seus dados
    - Semelhante ao modelo de difusão (p.e. televisão)
    - Oposto do modelo da telefonia
-

# Multicast: considerações básicas

---



- Qual é o problema ?
    - Necessário conhecer todos os receptores de um grupo *multicast*
    - Necessário criar uma árvore (*spanning tree*) de receptores
  - Considerações importantes:
    - Minimizar tráfego não desejado
    - Minimizar estado (sessões) nos *routers*
    - Escalabilidade quando o número de receptores (ou emissores) aumenta
    - Rapidez no estabelecimento de novas sessões
    - Robustez
-



- Grupos *multicast* sobrepõem-se a redes
  - Os destinatários de informação *multicast* podem estar em diferentes redes
  - Na mesma rede nem todos os computadores podem querer receber a informação *multicast*
- Como criar um grupo *multicast*?
- Quem pode ser adicionado ao grupo *multicast*?
- Como sair do grupo *multicast*?

# Encaminhamento *Multicast*

---



- *Debugging* é difícil
  - Protocolos e aplicações imaturas
  - Interoperabilidade entre protocolos *multicast* e *unicast*
  - Instabilidade no encaminhamento
  - O modelo *multicast* quebra os modelos de tarifação e de gestão das empresas telco/ISP
-



- Introdução
- Grupos Multicast: IGMP
- Encaminhamento Multicast
  - Distance Vector Multicast Routing Protocol
  - Core Base Trees
  - Protocol Independent Multicast
  - Border Gateway Management Protocol



- Endereços de Grupos
  - 224.0.0.0 – 239.255.255.255
- Endereços “Classe D” – bits de maior peso “1110”
- Endereços de Grupos reservados:
  - 224.0.0.0 – 224.0.0.255:
    - 224.0.0.1 – Todos os *hosts* nesta subrede
    - 224.0.0.2 – Todos os *routers* nesta subrede
    - 224.0.0.5 – Todos os *routers* OSPF
    - 224.0.0.6 – Todos os *routers* OSPF que sejam DR
    - 224.0.0.9 – Todos os *routers* RIPv2



# Reserva de endereços multicast

---



- Problema
    - Endereços multicast são um recurso limitado
    - Esquema actual de reserva de endereços multicast não escala e torna o encaminhamento multicast mais difícil
  - Solução
    - Usar endereços reservados dinamicamente
    - Reserva de endereços determina a raiz da árvore partilhada
    - Reserva de endereços hierárquica escala melhor e ajuda o encaminhamento multicast
-

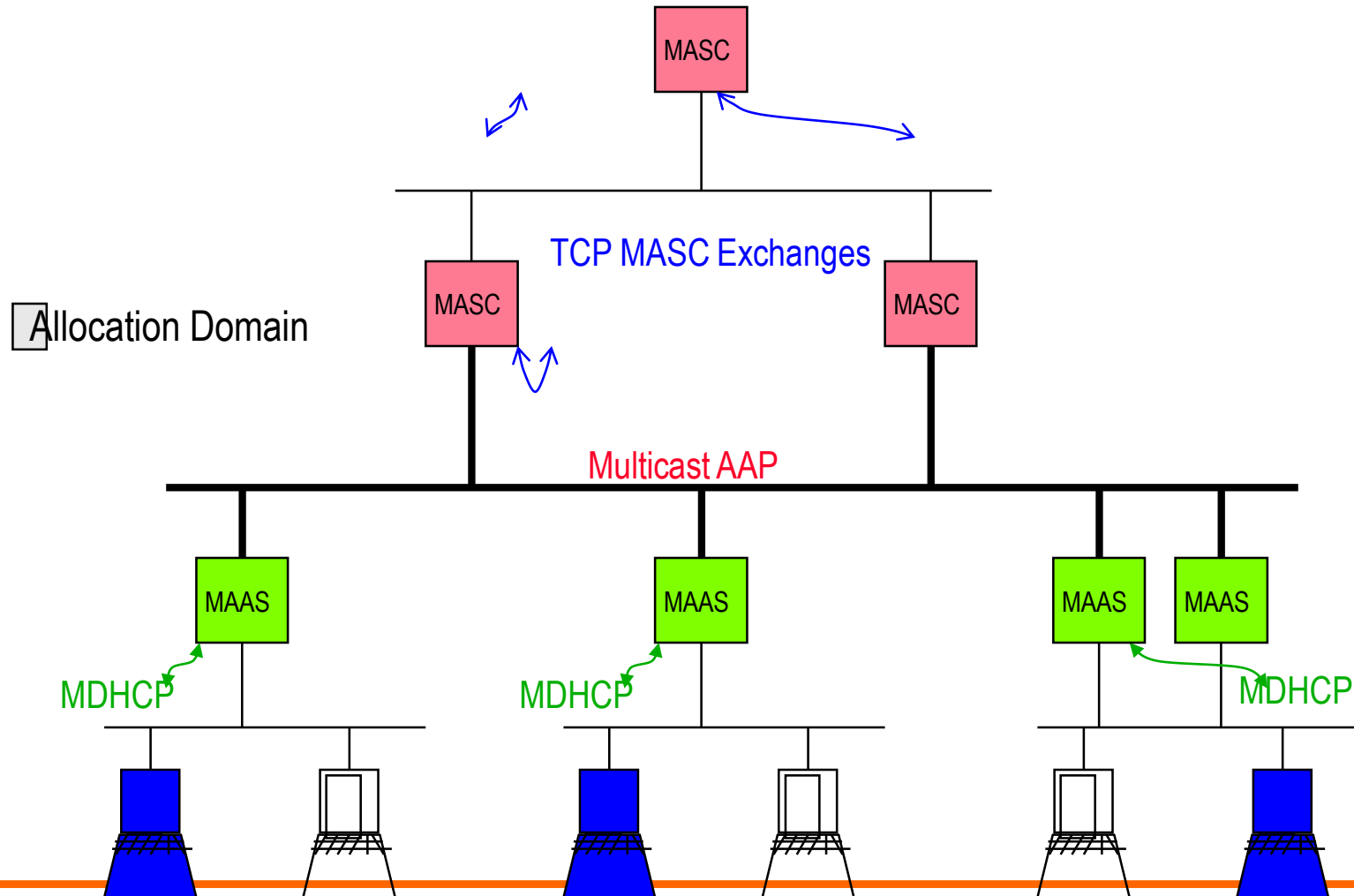
# Arquitectura de reserva de endereços multicast

---

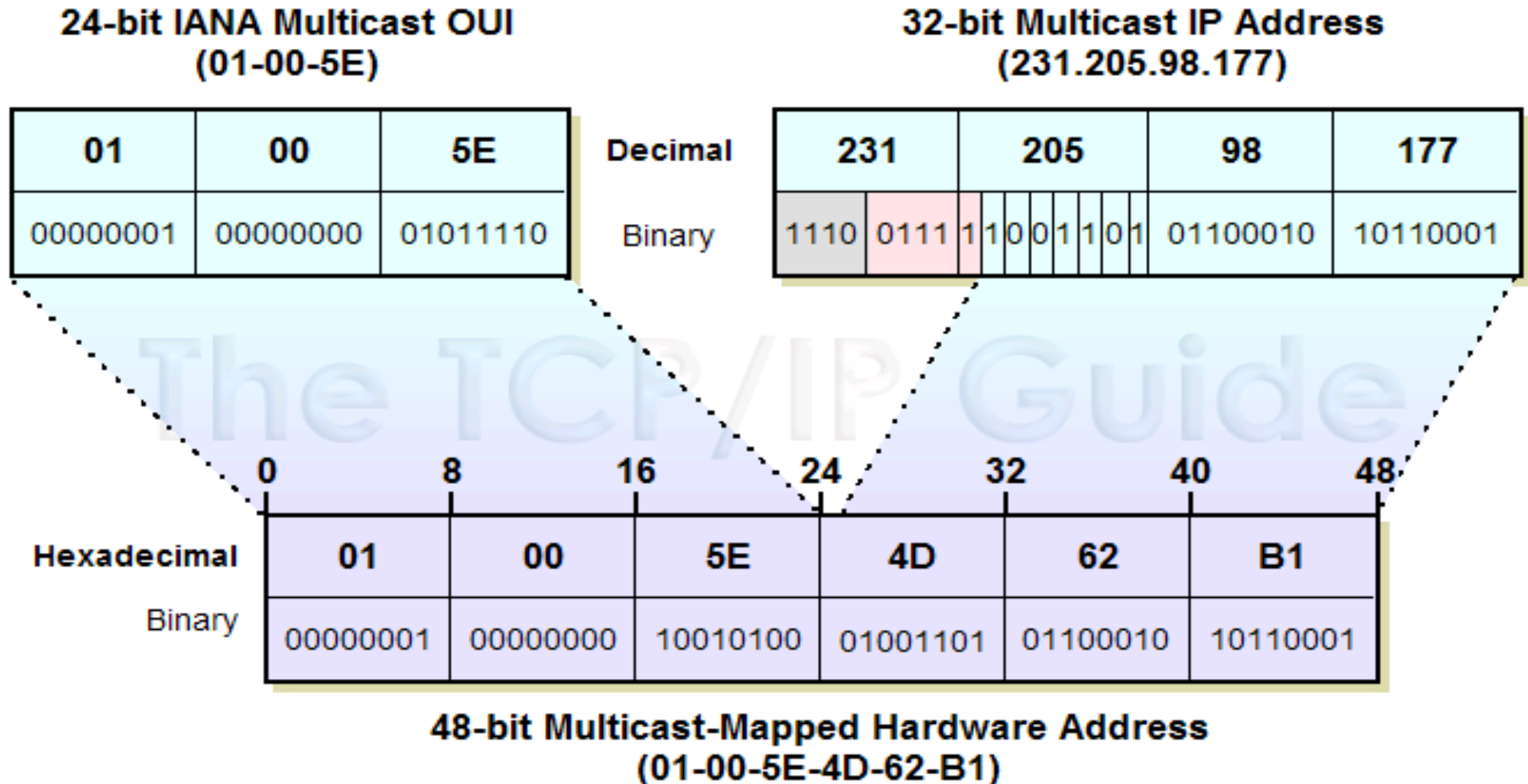


- Multicast Address Set Claim (MASC)
  - Protocolo para reservar endereços multicast atribuídos a domínios
  - Algoritmo: Ouvir e detectar a ocorrência de colisões
- Address Allocation Protocol (AAP)
  - Protocolo para reservar endereços multicast dentro dos domínios
  - Usado pelos servidores Multicast Address Allocation Servers (MAAS)
- MDCHP (Multicast DHCP)
  - Protocolo para os terminais pedirem endereços multicast
  - Extensão ao DHCP (Dynamic Host Configuration Protocol)

# Multicast Address Allocation Example



# Conversão Endereços – IP Multicast - MAC



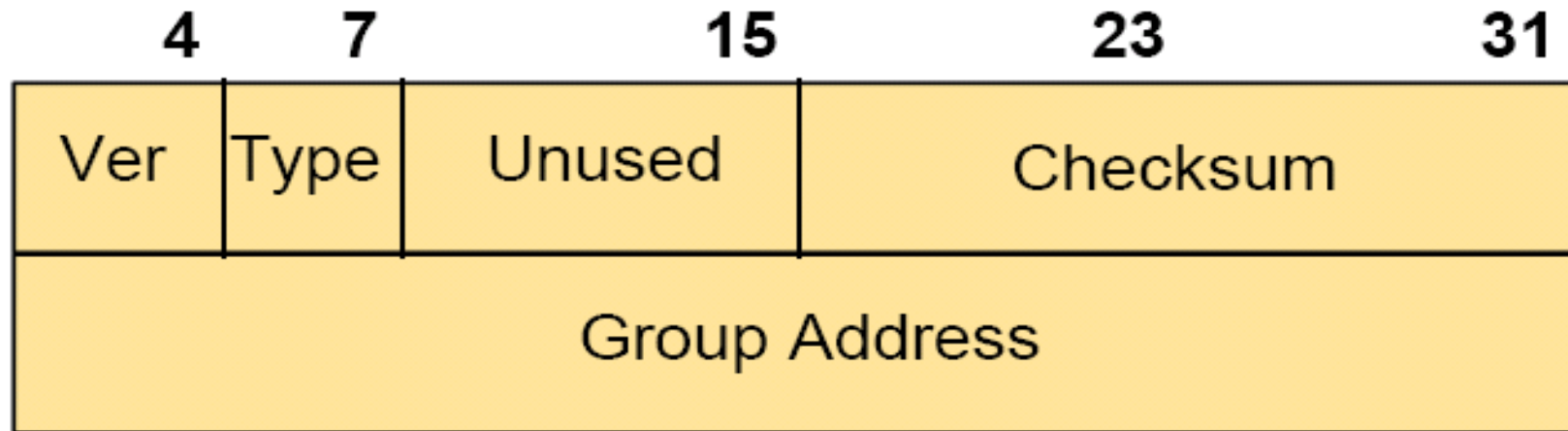


- Os *hosts* informam os *routers* quais os grupos que estão interessados em receber
- Os *routers* solicitam informação sobre quais os grupos que os *hosts* ainda estão interessados em receber
- IGMPv1 definido no RFC 1112
- IGMPv2 definido no RFC 2236
- IGMPv3 ainda em *draft*



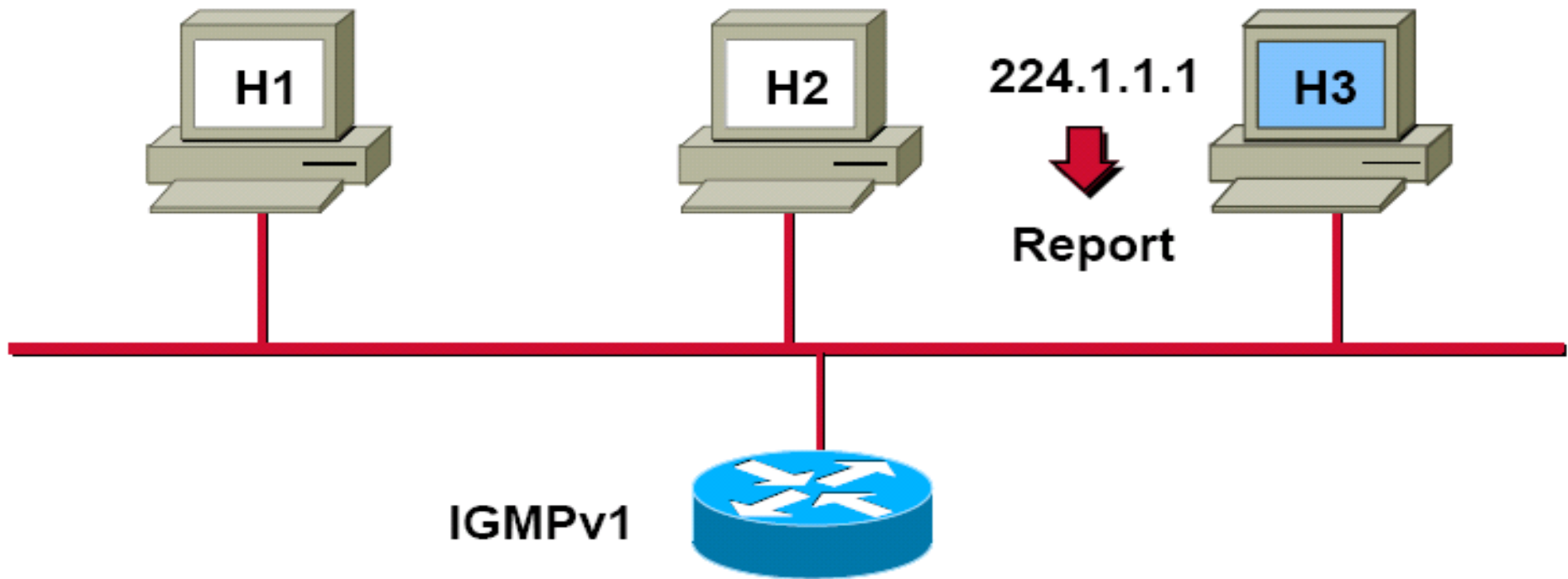
- Define 2 tipos de mensagem
- *Membership Queries*
  - O *querier* envia uma mensagem *query* de IGMP para o endereço 224.0.0.2 com ttl=1
  - Apenas um *router* por LAN está encarregue de enviar *queries*
  - O intervalo dos *queries* varia entre 60-120 segundos (configurável)
- *Membership Reports*
  - A mensagens IGMP de *report* enviada por um *host* suprime as dos outros *hosts*
  - Só existe um *report* por LAN
  - Existem *reports* não solicitados – são enviados quando um *host* quer receber um grupo

# IGMPv1 – Formato do Pacote



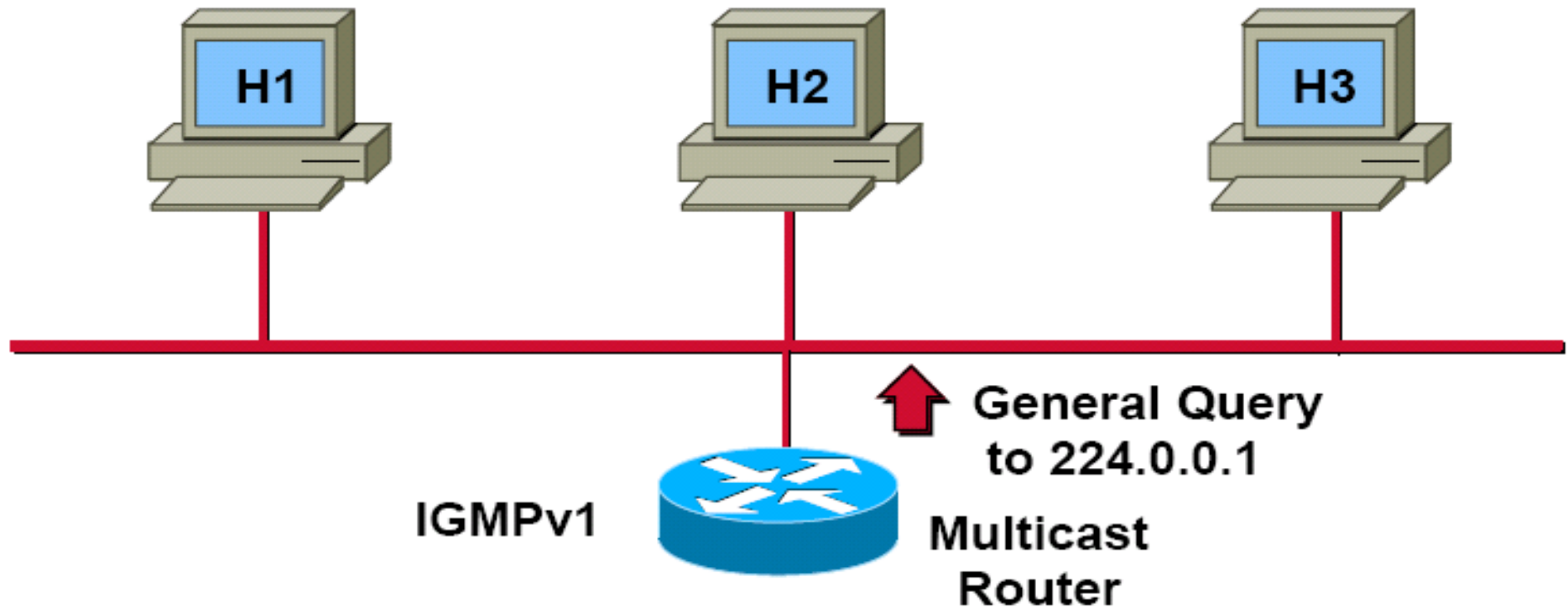
- Ver – versão IGMP
- Tipo
  - 1 – Query
  - 2 – Report
- Group address – Endereço do grupo *multicast*

# IGMPv1 – Juntar-se a um grupo

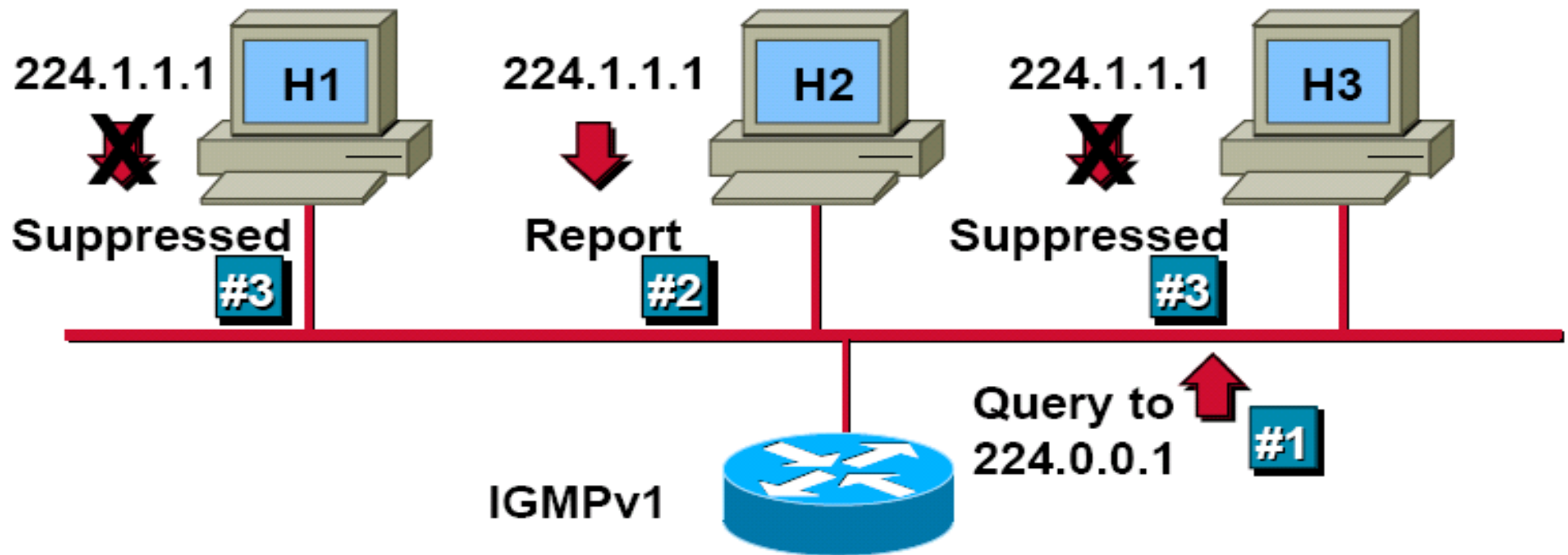




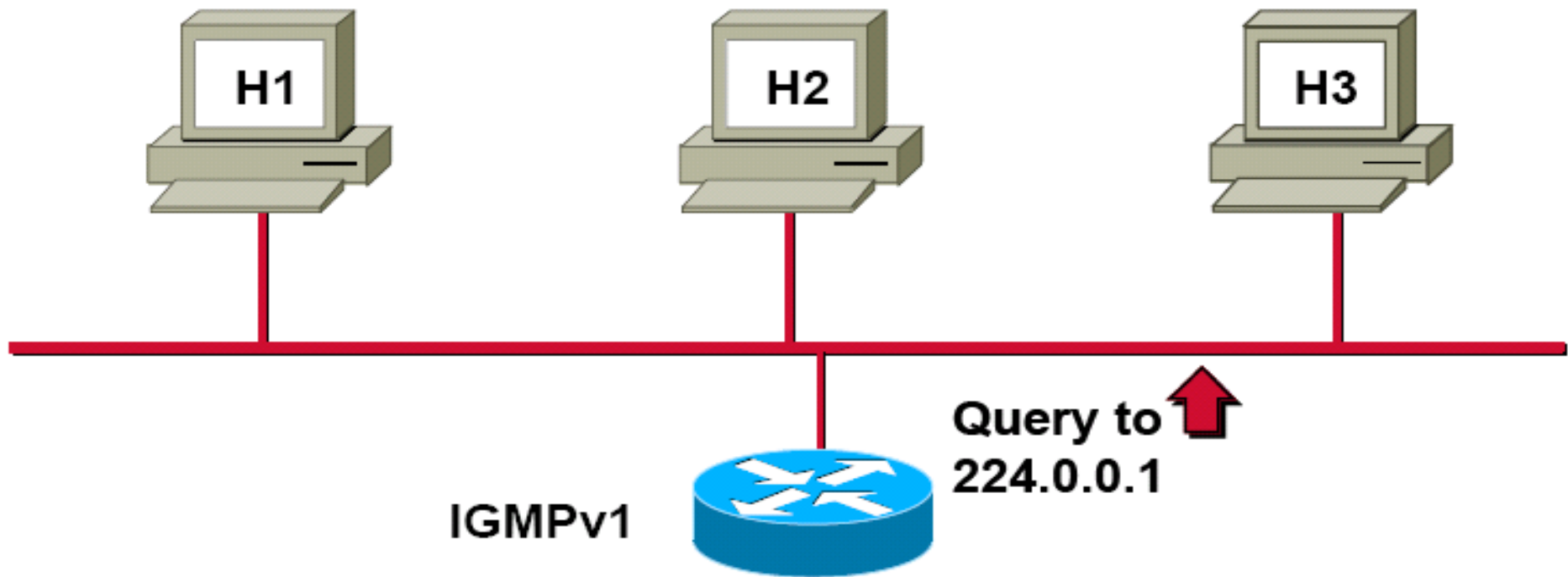
# IGMPv1 – General Queries



# IGMPv1 – Manutenção de um grupo



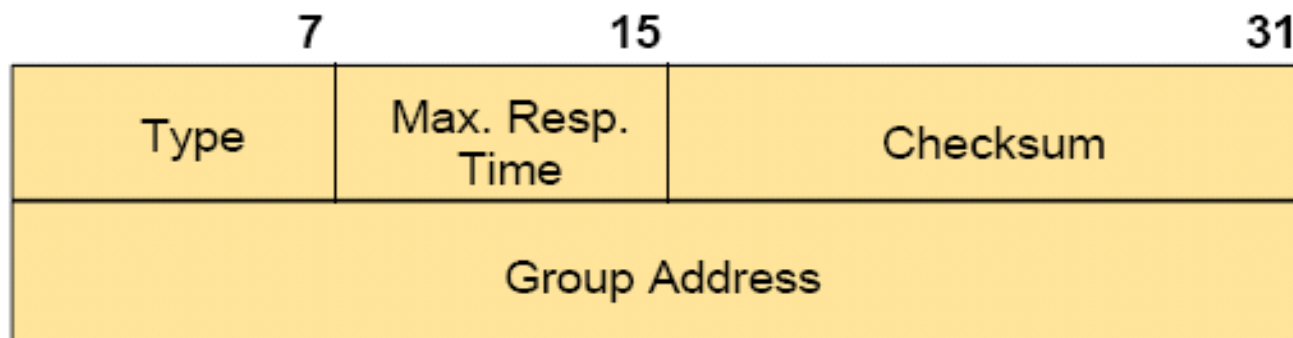
# IGMPv1 – Sair de um grupo





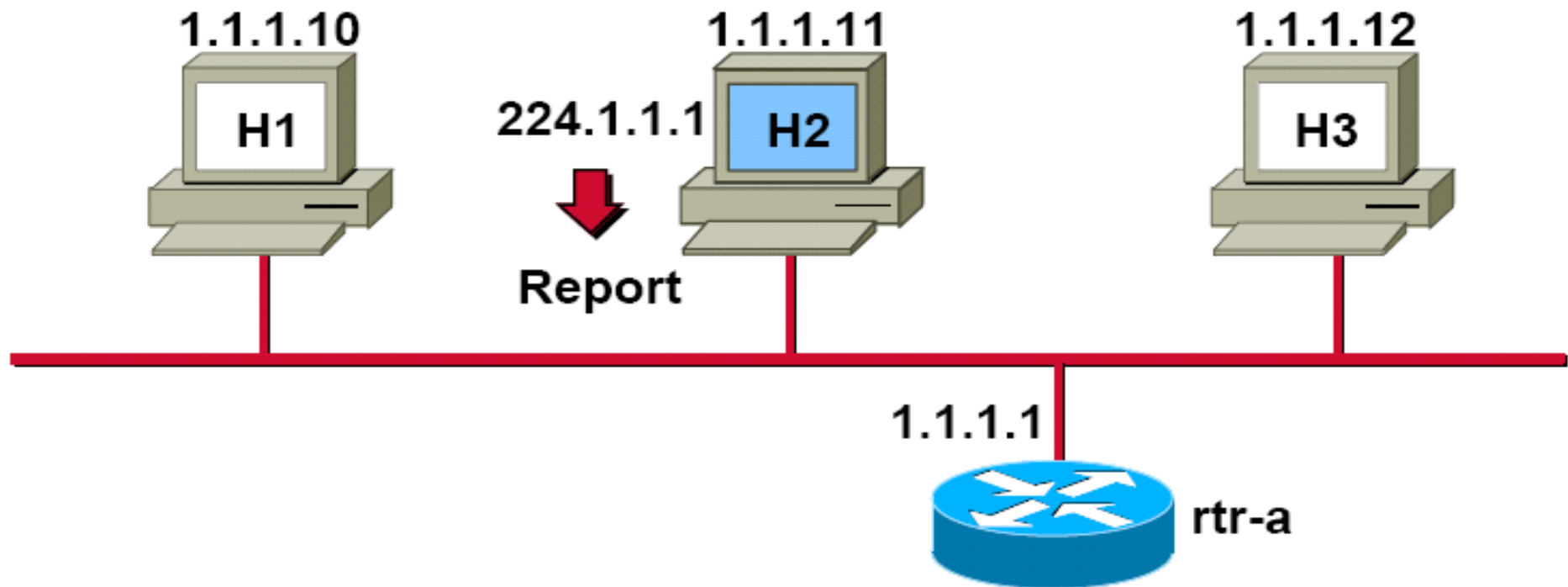
- RFC 2236 – define mais duas mensagens e um mecanismo de eleição do *querier*
  - O *router* com o endereço IP mais baixo é eleito o *querier*
- *Group-specific query*
  - Permite ao *router* saber se existe alguém interessado num grupo específico
- *Leave group*
  - O *host* envia uma mensagem de *leave* quando deixa de estar interessado num grupo e é o último membro

# IGMPv2 – Formato do Pacote

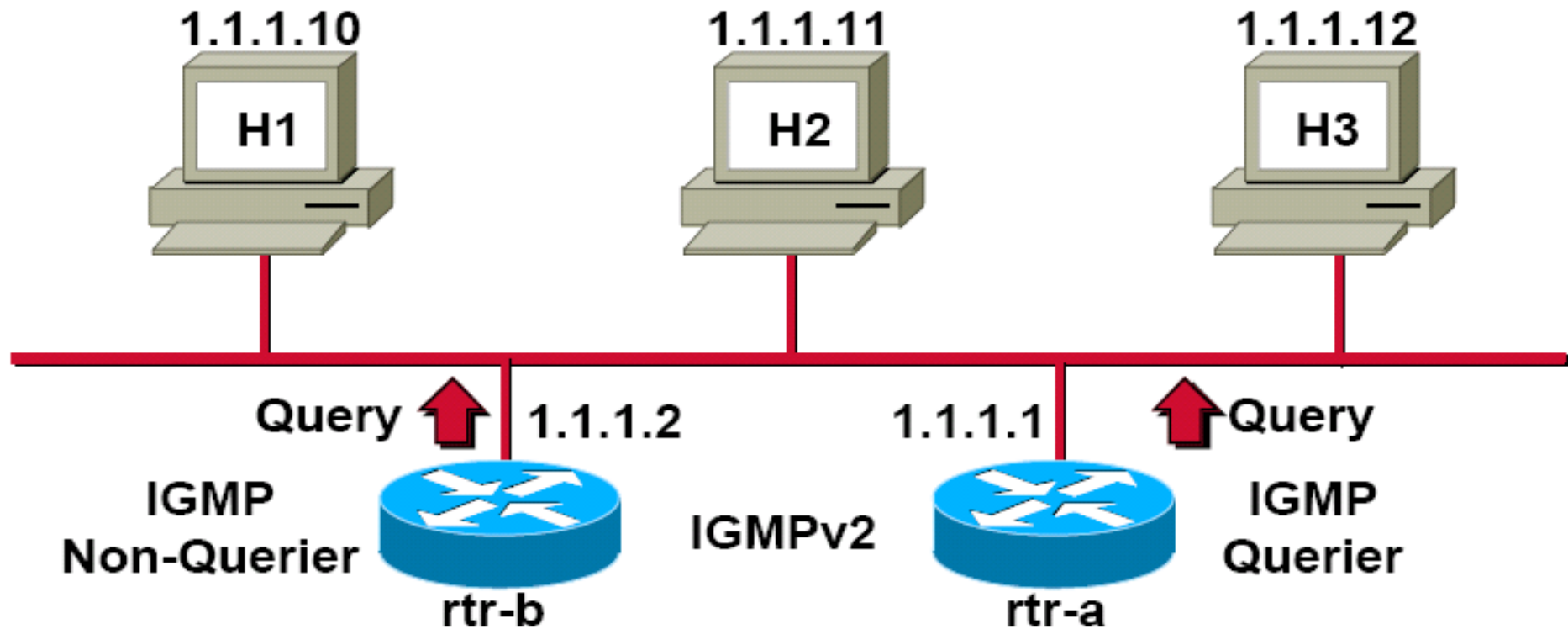


- Tipo
  - 0x11 – Query
  - 0x12 – Report v1
  - 0x16 – Report v2
  - 0x17 – Leave Group
- Max. Resp. Time – Tempo máximo para enviar uma resposta - *report*
- Group address – Endereço do grupo *multicast* (0.0.0.0 para todos os grupos – *general query*)

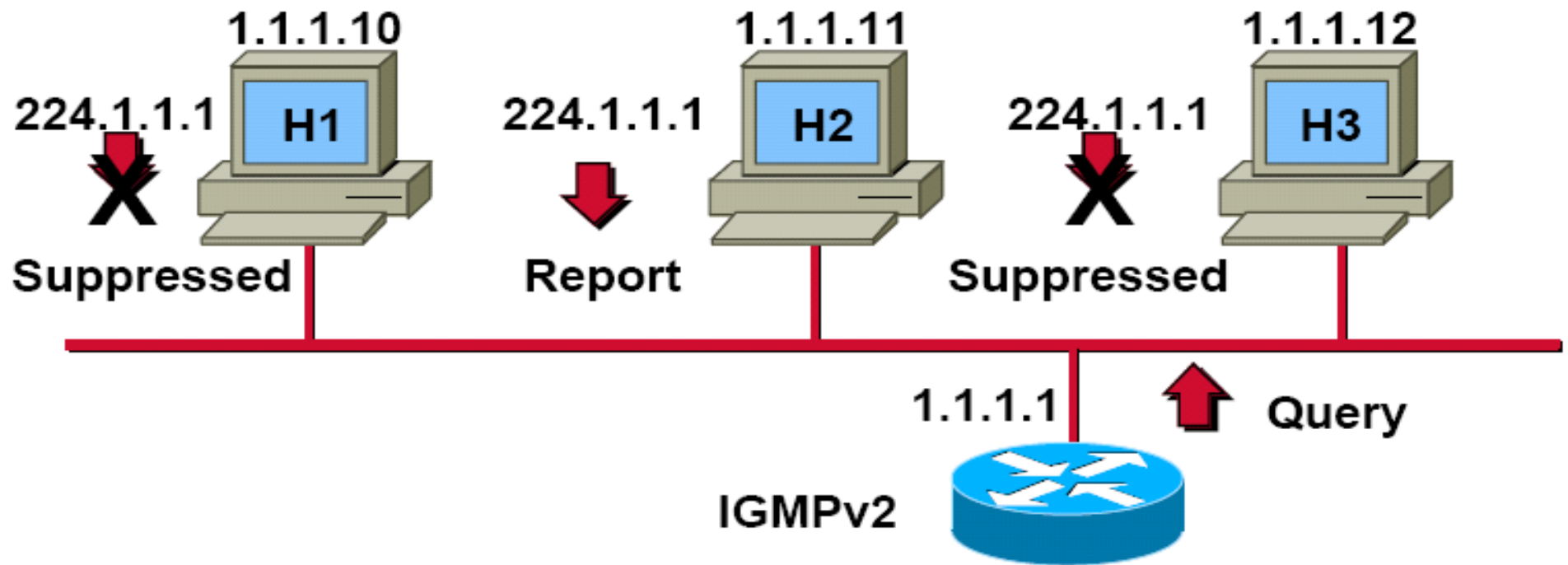
# IGMPv2– Juntar-se a um grupo



# IGMPv2 – Eleição do Querier

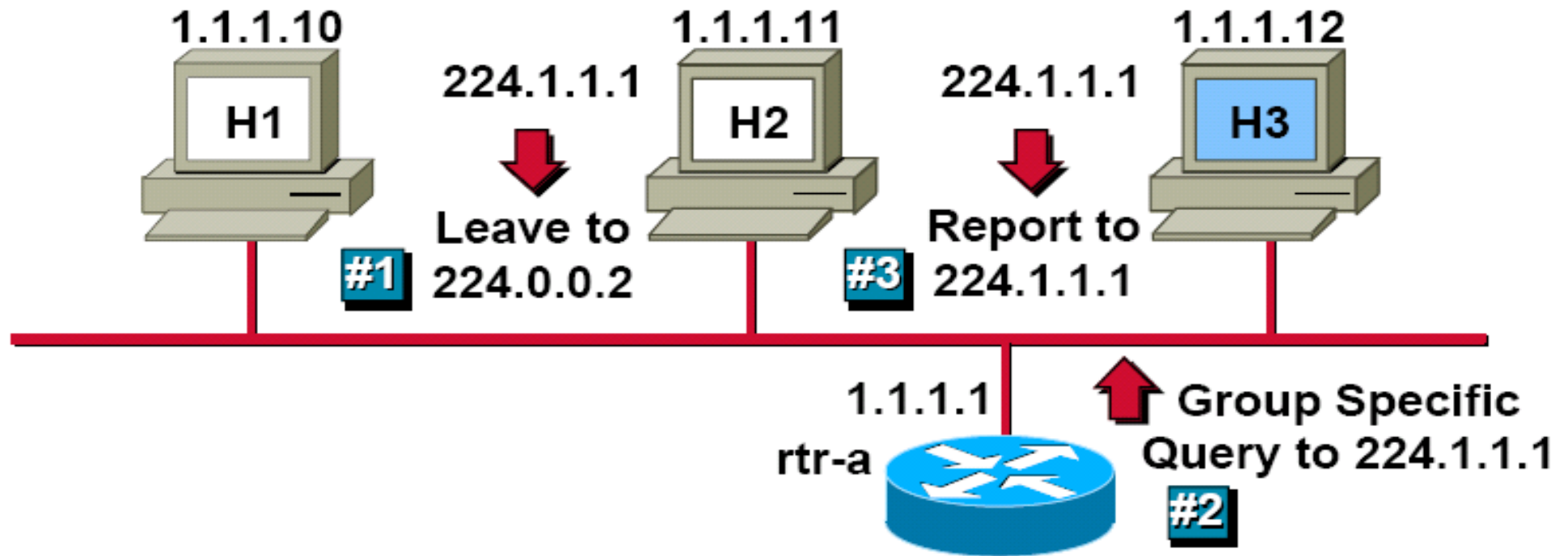


# IGMPv2 – Manutenção de um grupo

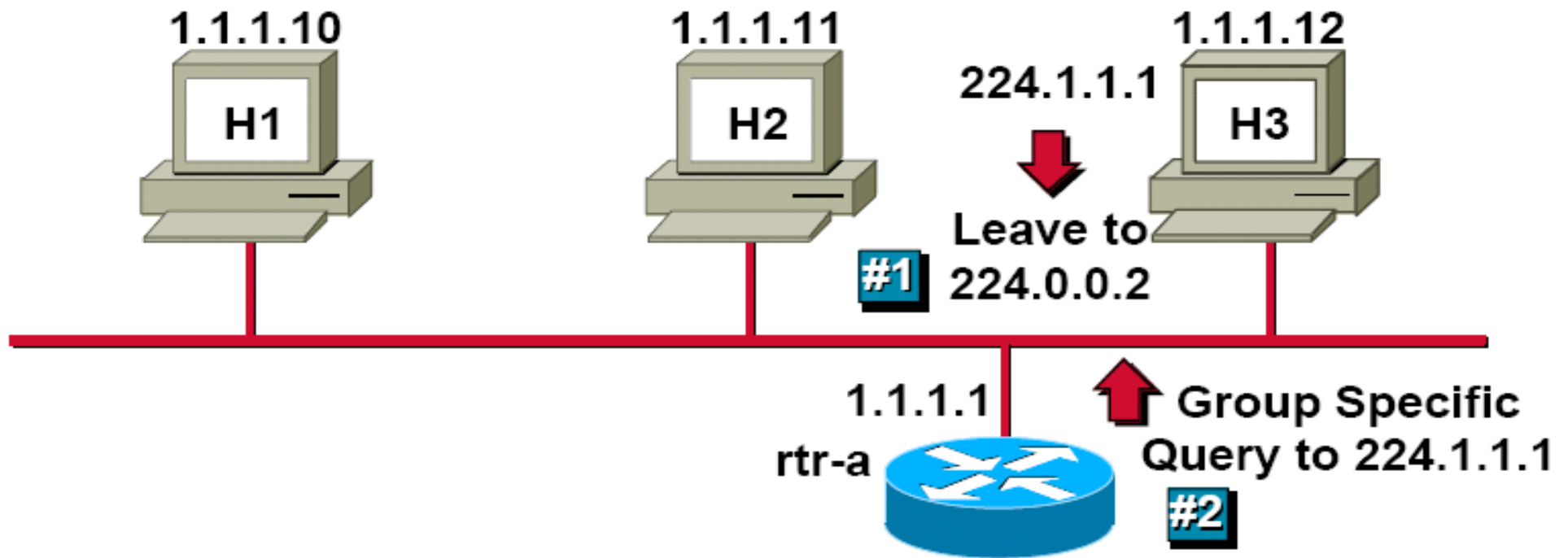




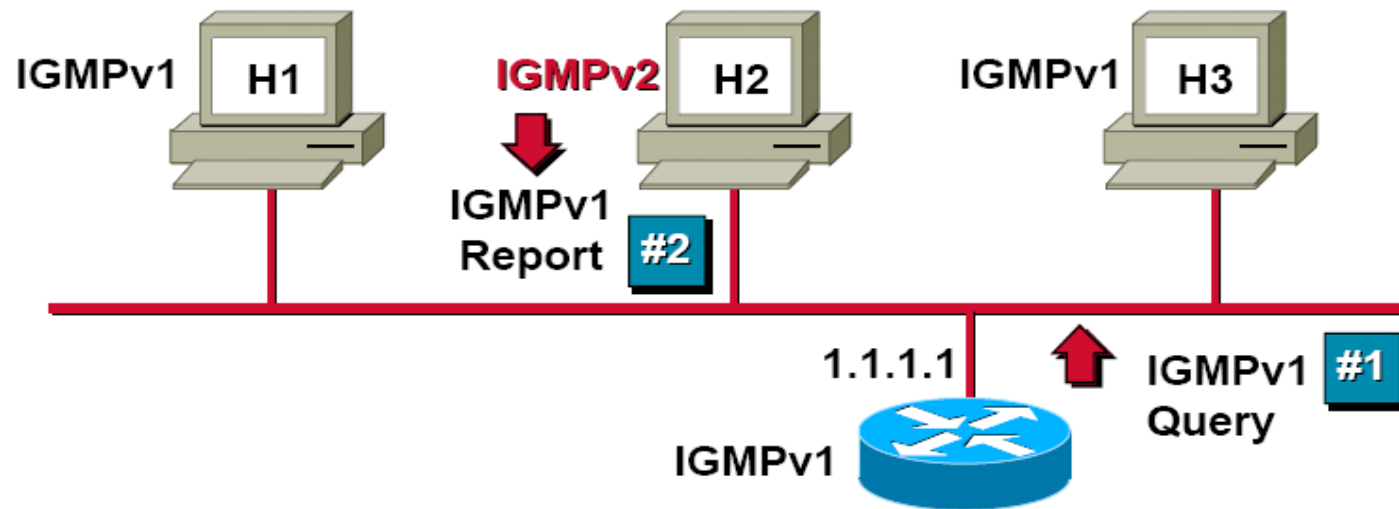
# IGMPv2 – Sair de um grupo



# IGMPv2 – Sair de um grupo (cont.)

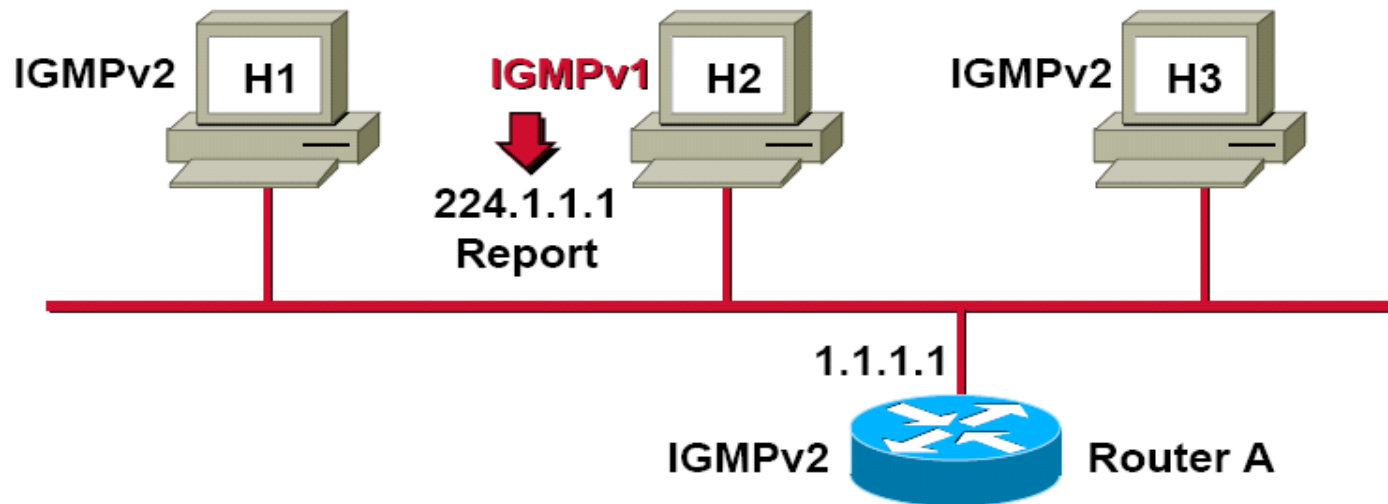


# Interoperabilidade IGMPv1 – IGMPv2



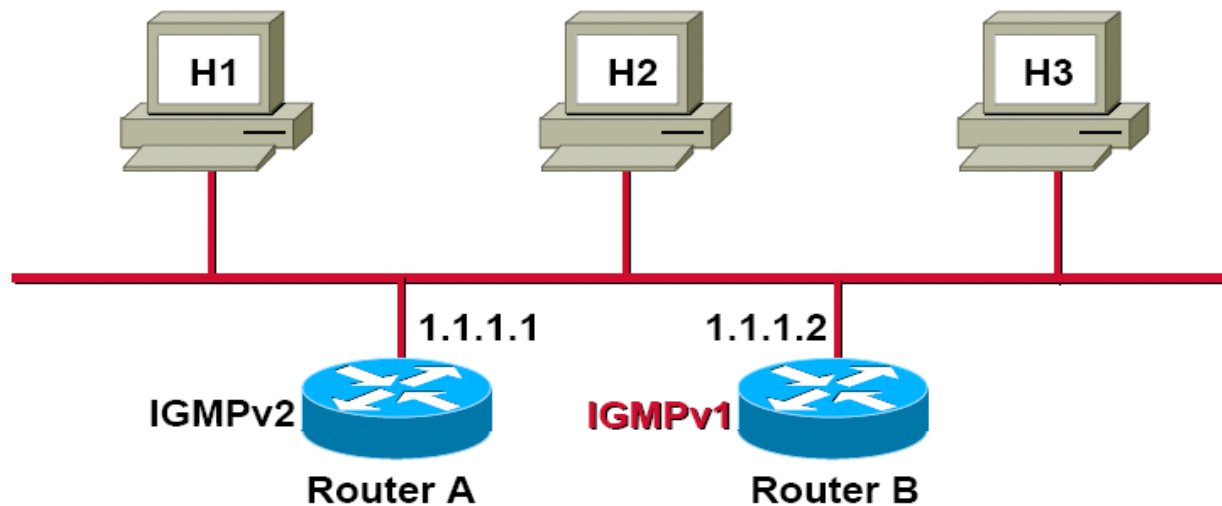
- H2 – Tem de enviar sempre *reports* IGMPv1
- H2 – Pode suprimir *leaves* IGMPv2

# Interoperabilidade IGMPv1 – IGMPv2



- Router A – Tem de perceber que existe um membro em IGMPv1 para o grupo 224.1.1.1
- Router A – Tem de ignorar todos os *leaves* para o grupo 224.1.1.1

# Interoperabilidade IGMPv1 – IGMPv2

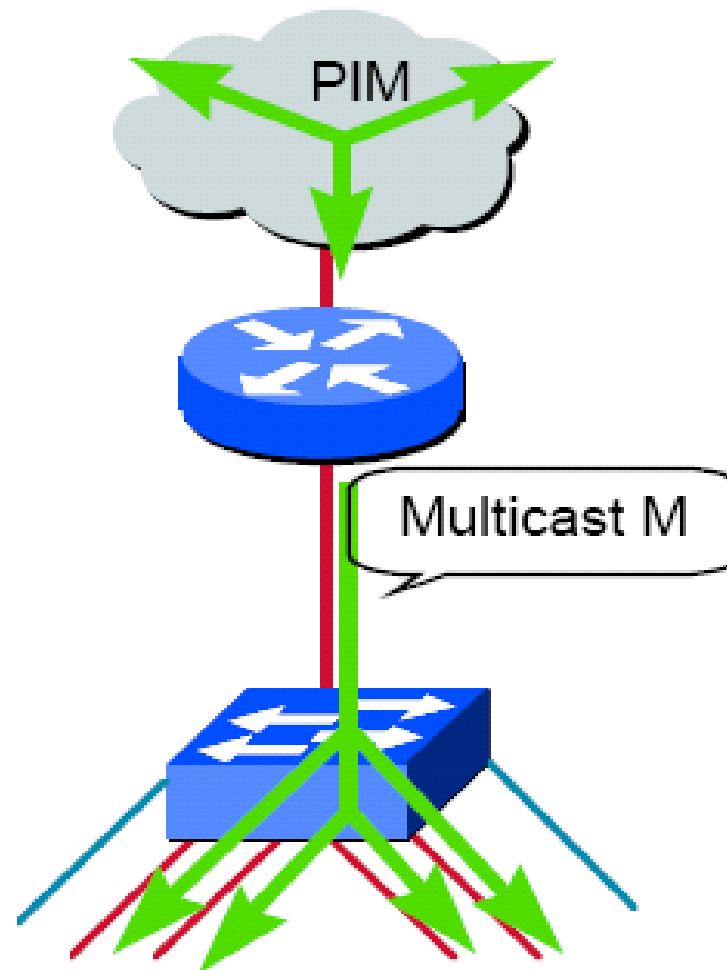


- Router A – Tem de ser manualmente configurado para funcionar em IGMPv1

# Multicast na camada 2



- Problema: *Flooding* de tramas multicast para todas as portas de um *switch*
- Solução: Os *switch* têm de ficar também atentos ao IGMP e perceber que portas precisam de determinado grupo *multicast*





- Introdução
- Grupos Multicast: IGMP
- Encaminhamento Multicast
  - Distance Vector Multicast Routing Protocol
  - Core Base Trees
  - Protocol Independent Multicast
  - Border Gateway Management Protocol

# Encaminhamento Multicast: considerações básicas

---



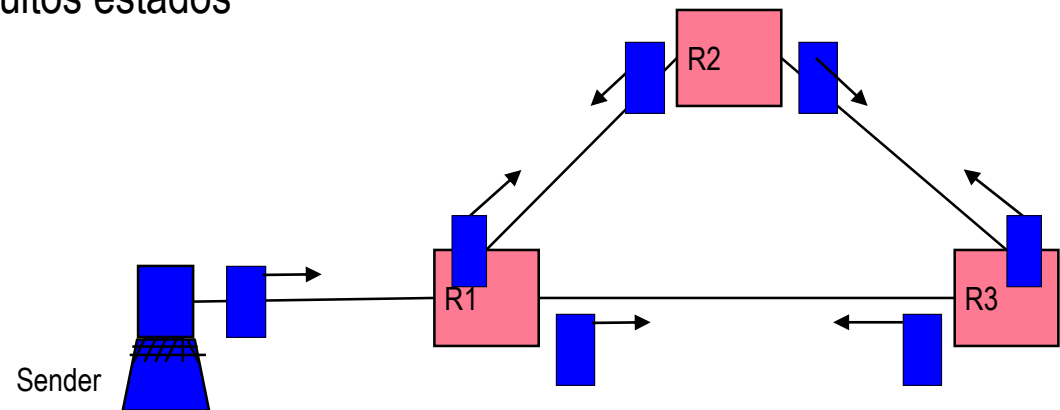
- Flood and prune
    - Distance Vector Multicast Routing Protocol (DVMRP)
    - Protocol Independent Multicast Dense Mode (PIM-DM)
  - Subscrição explícita (join)
    - Core Based Trees (CBT)
    - Protocol Independent Multicast Sparse Mode (PIM-SM)
  - A topologia da rede é conhecida por todos os routers
    - Protocolo de estado da ligação
    - MOSPF
-



# Protocolos do tipo flooding



- Enviar dados para todos os routers da rede
- Problemas graves:
  - Necessário evitar ciclos
  - Necessário enviar apenas uma vez para todos os nós de uma rede
  - Pode-se manter estado de cada pacote e verificar se este já visitou um determinado router anteriormente,
    - torna-se necessário manter muitos estados



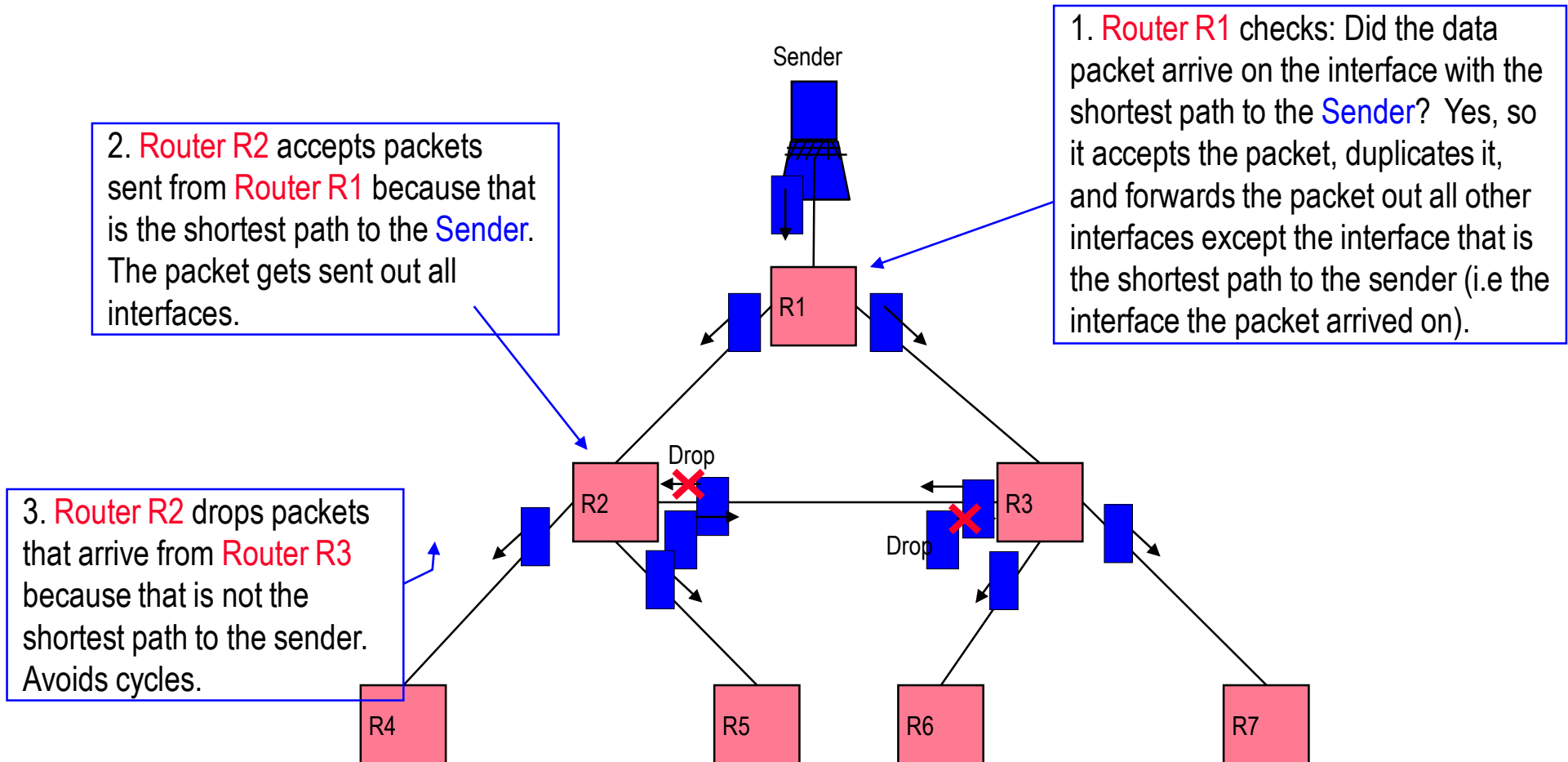
# Reverse Path Forwarding (RPF)

---



- Técnica simples para construir árvores:
    - Em encaminhamento unicast, os routers enviam para o destino pelo caminho mais curto
    - Em encaminhamento multicast, os routers enviam para todas as interfaces excepto por aquela que têm o caminho mais curto para a origem
-

# Reverse Path Forwarding Example



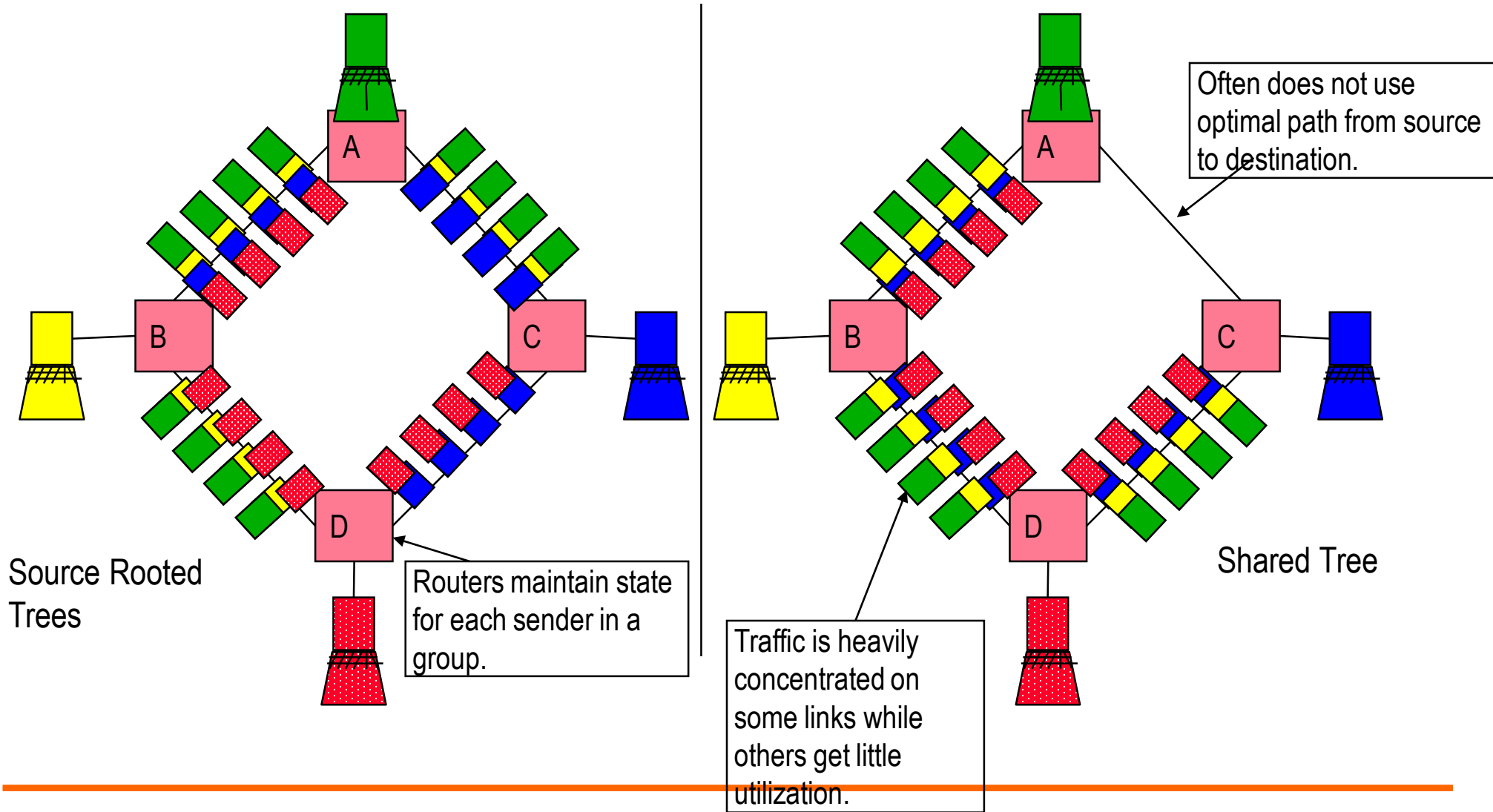
# Escolhas na distribuição de dados

---



- Árvores centradas na origem (source rooted trees)
    - É mantido estado nos routers para cada emissor
    - Constrói árvore SPT de cada emissor para os receptores
    - Atrasos mínimos das origens para os destinos
  - Árvores partilhadas na origem (shared trees)
    - Todos os emissores pertecem à mesma árvore SPT de distribuição
    - É mantido estado nos routers apenas para os grupos multicast desejados
    - Não é mantido estado por emissor (excepto para IGMPv3)
    - Maior atraso na distribuição de dados
-

# Árvores centradas na origem vs. Árvores partilhadas





- Introdução
- Grupos Multicast: IGMP
- Encaminhamento Multicast
  - Distance Vector Multicast Routing Protocol
  - Core Base Trees
  - Protocol Independent Multicast
  - Border Gateway Management Protocol

# Distance Vector Multicast Routing (DVMRP)



- Árvore SPT (spanning tree)
  - Constrói *shortest path tree*
  - Mínimo número de saltos da origem para os receptores
- Estende os protocolos de encaminhamento unicast do tipo distance vector
- Algoritmo do tipo *flood and prune* :
  - No início os dados são enviados para todos os routers da rede usando o RPF
  - Mensagens do tipo *prune* são enviadas pelo routers para grupos não desejados para se limitar a dimensão da rede
  - É necessário manter estado nos routers para todos os grupos não desejados
  - Mensagens “prune” são enviadas periodicamente devido à ocorrência de timeout (um router que enviou um prune pode agora desejar pertencer a esse grupo)



- Truncated Reverse Path Muticast
  - Versão otimizada do Reverse Path Forwarding (RPF)
  - Truncating: nenhuns pacotes são enviados para redes sem receptores interessados nesse grupo
- Pruning:
  - Mensagens do tipo prune são enviadas se não houver receptores subscritos nesse grupo
  - Necessário manter estado para cada grupo não desejado
- Grafting:
  - Sempre que houver uma subscrição para um grupo cancelado (através de uma mensagem de prune) é necessário enviar uma mensagem graft para cancelar a mensagem prune enviada periodicamente
  - Na ocorrência de um join (IGMP) ou graft (de outro router), remove-se o estado prune e propaga-se uma mensagem graft

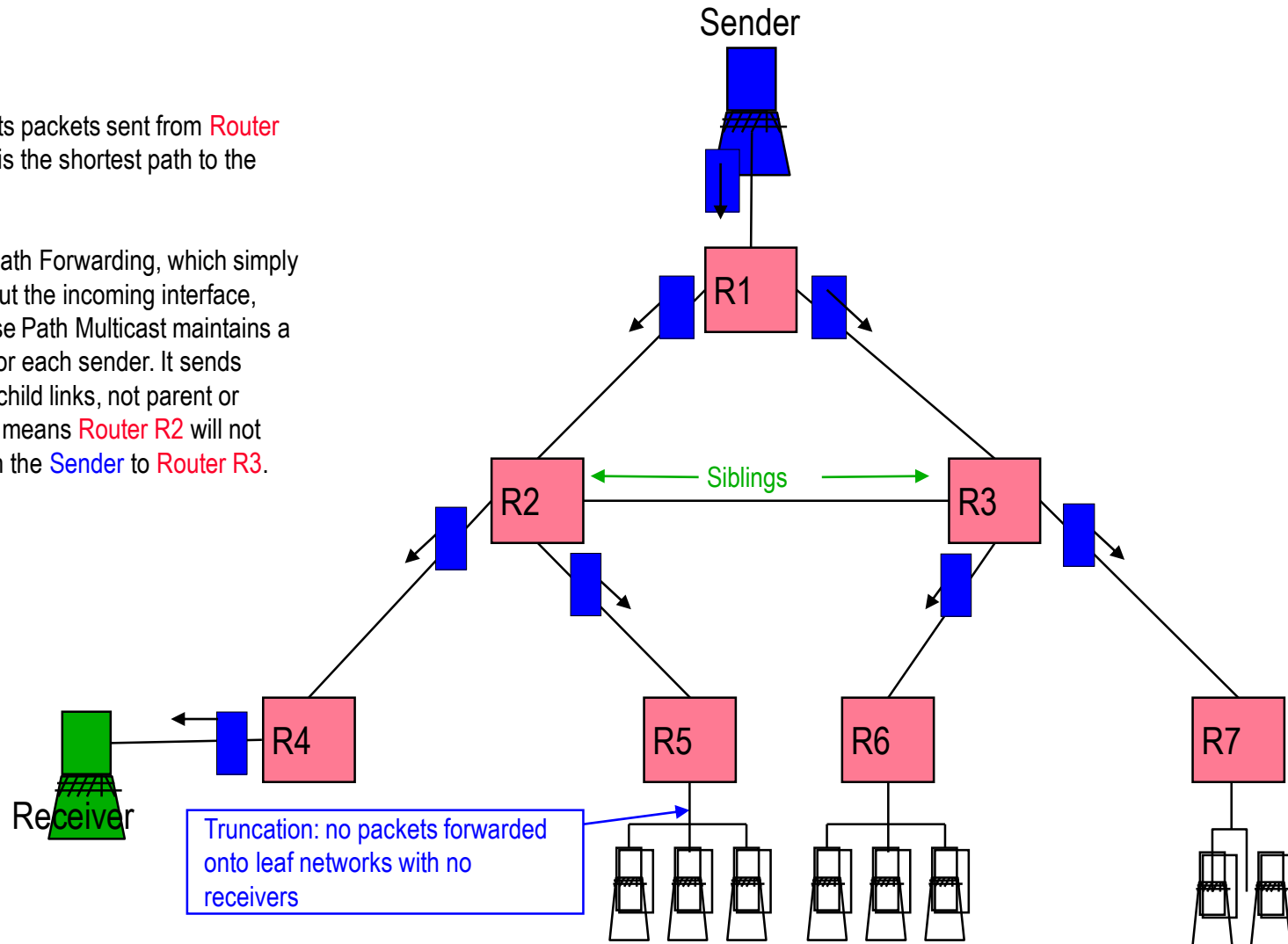


# Truncated Reverse Path Multicast Example

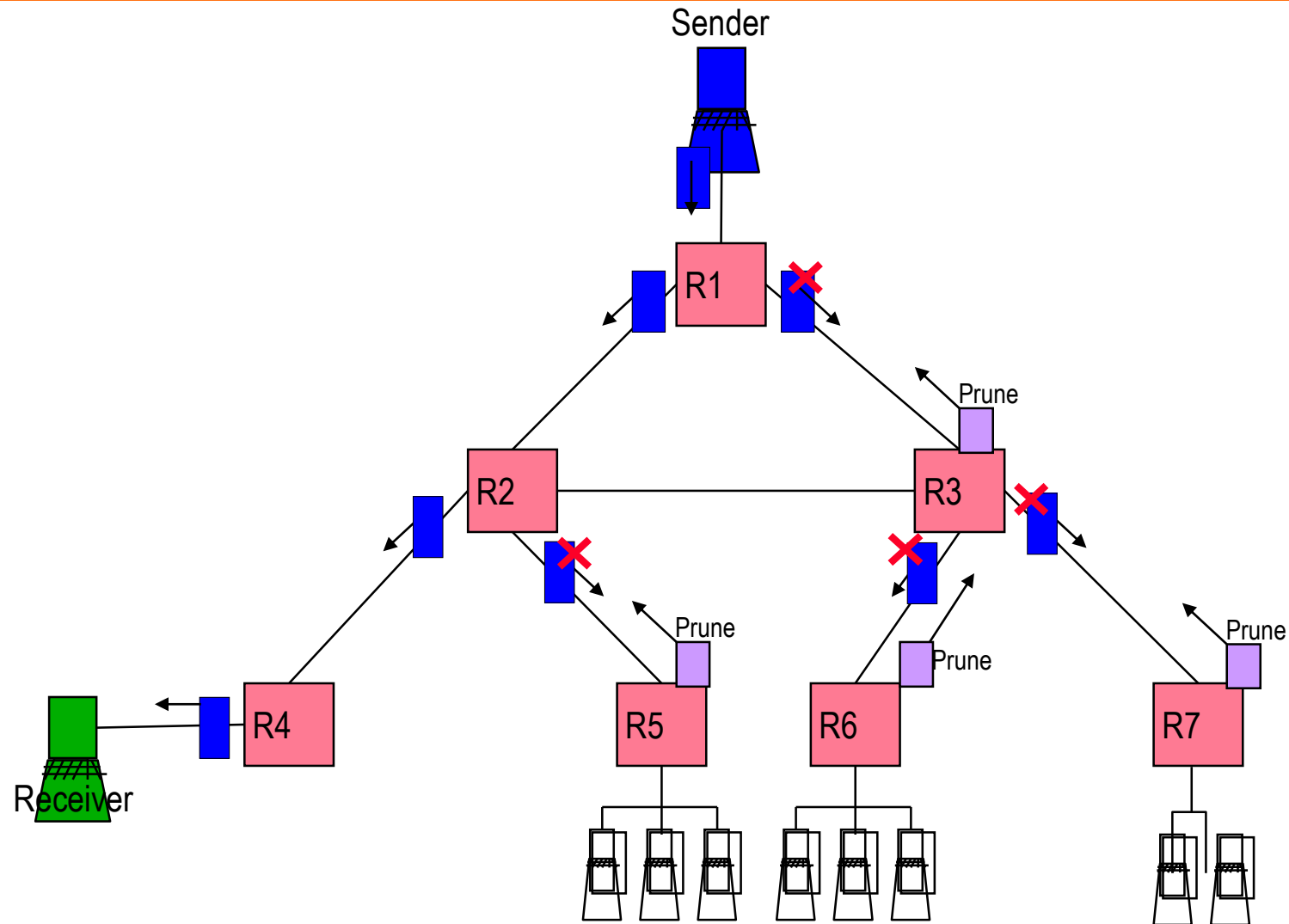


Router R2 accepts packets sent from Router R1 because that is the shortest path to the Sender.

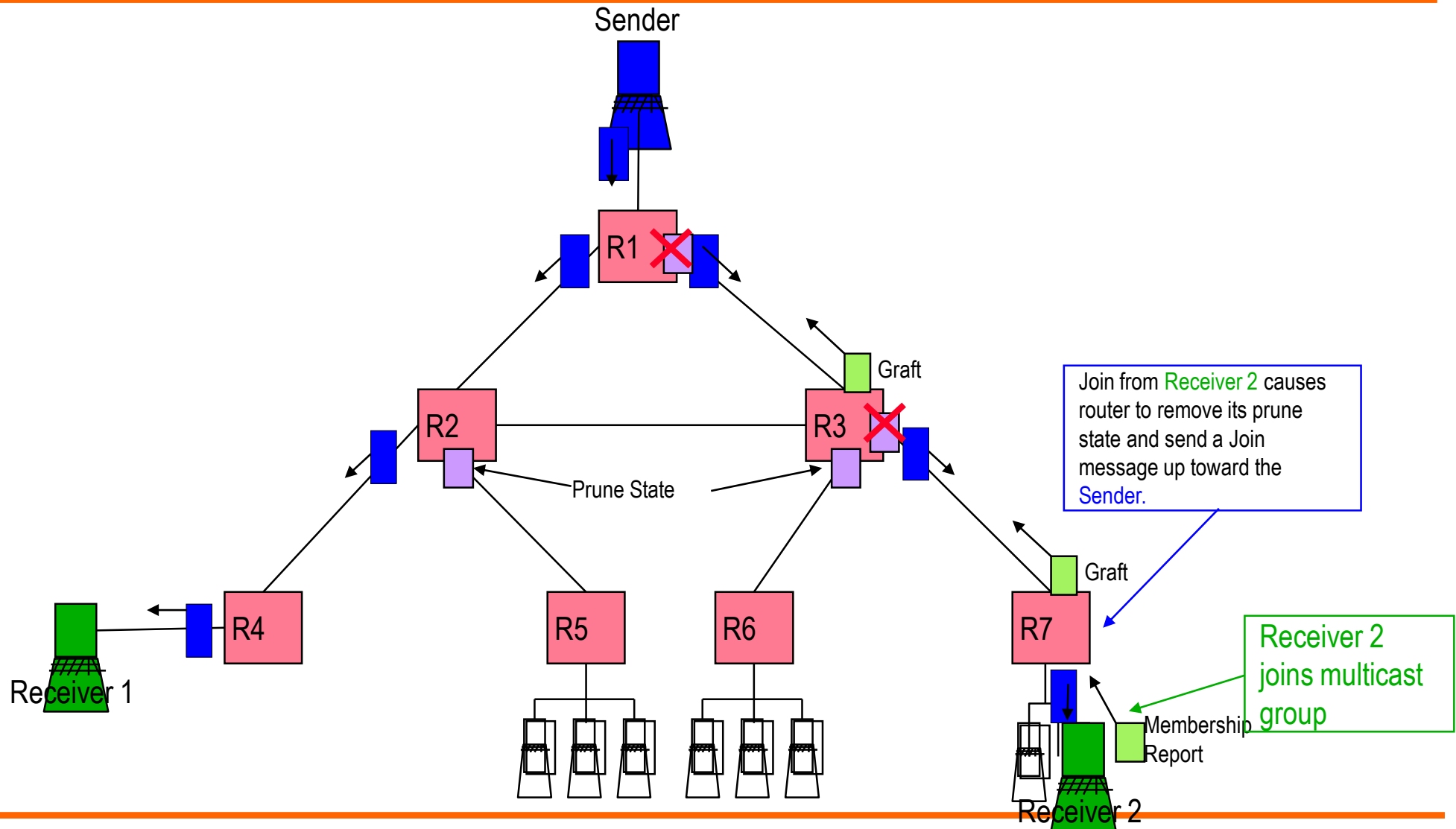
Unlike Reverse Path Forwarding, which simply forwards out all but the incoming interface, DVMRP's Reverse Path Multicast maintains a list of child links for each sender. It sends packets only out child links, not parent or sibling links. This means Router R2 will not forward data from the Sender to Router R3.



# DVMRP Pruning Example



# DVMRP Grafting Example



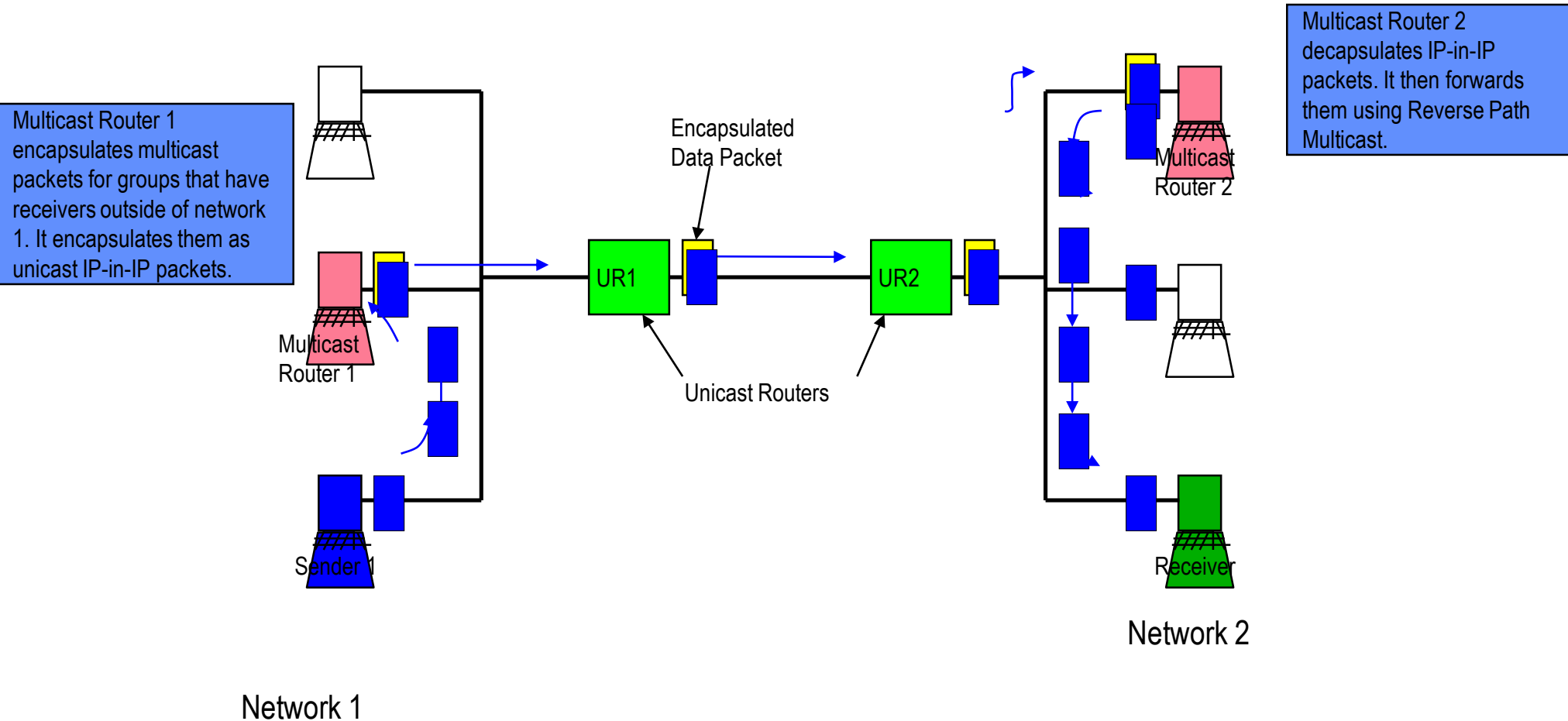
# Multicast Tunneling

---



- Problema:
    - Nem todos os routers são capazes de efectuar *multicast*
    - No entanto, pode-se querer interligar domínios entre routers não multicast entre eles
  - Solução:
    - Encapsular pacotes multicast em pacotes unicast
    - Tunel de tráfego multicast entre routers não multicast
  - MBONE:
    - Rede multicast virtual, subconjunto da internet
    - Regiões nativas multicast são interligadas com túneis
-

# Exemplo Multicast Tunneling



# Problemas do DVMRP

---



- Alguns dos problemas dos protocolos distance vector mantêm-se
  - É preciso manter estado para grupos não desejados
  - Uso de largura de banda é intenso
    - Data flooding periódico por grupo
      - Não existem joins explícitos, e o estado prune faz timeout
    - Não é adequado a redes heterogéneas (com protocolos de encaminhamento unicast diferentes)
  - Escala muito mal quando existe um número elevado de emissores
-



- Introdução
- Grupos Multicast: IGMP
- Encaminhamento Multicast
  - Distance Vector Multicast Routing Protocol
  - Core Base Trees
  - Protocol Independent Multicast
  - Border Gateway Management Protocol

# Core Based Trees (CBT)

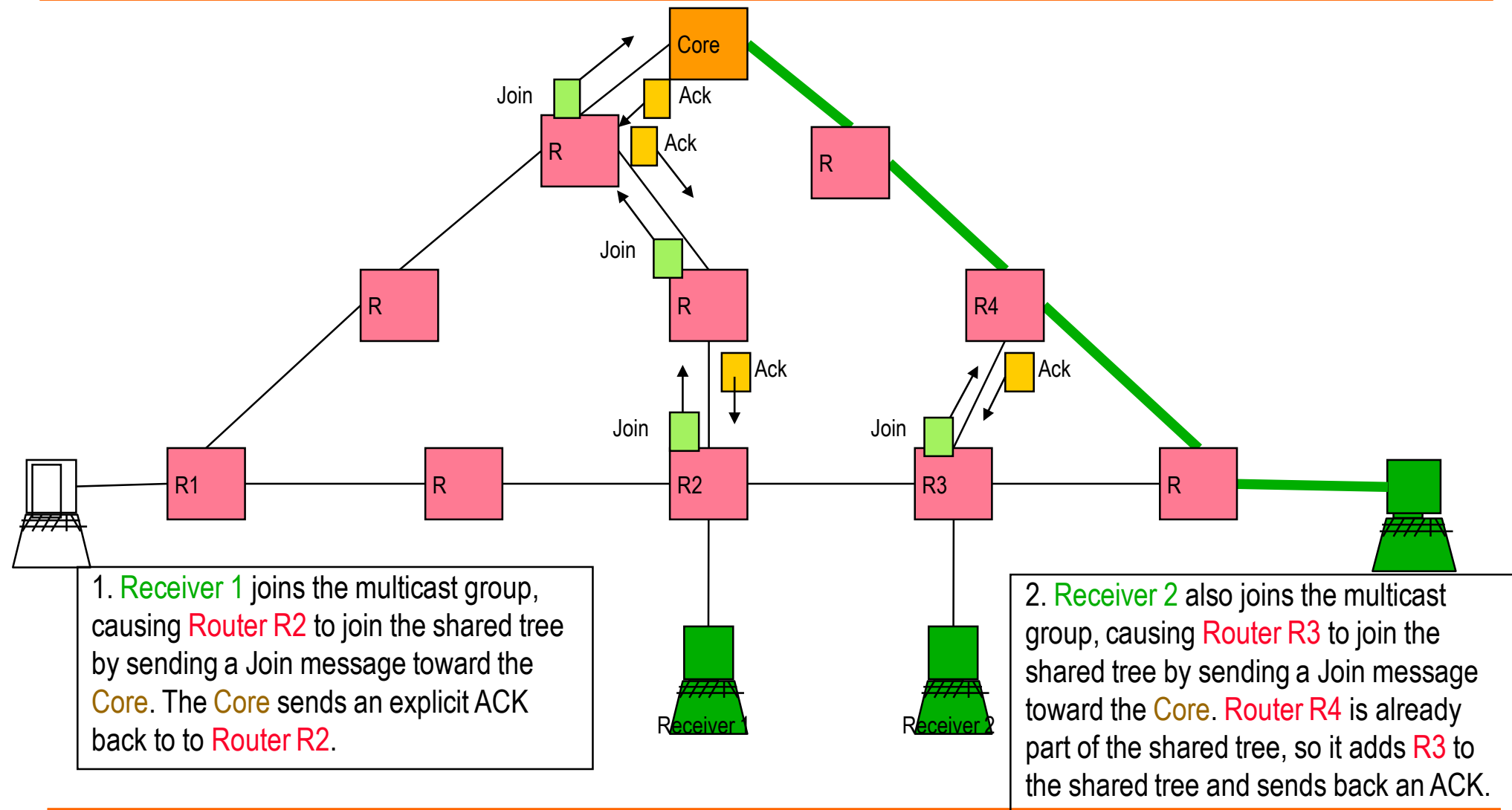
---



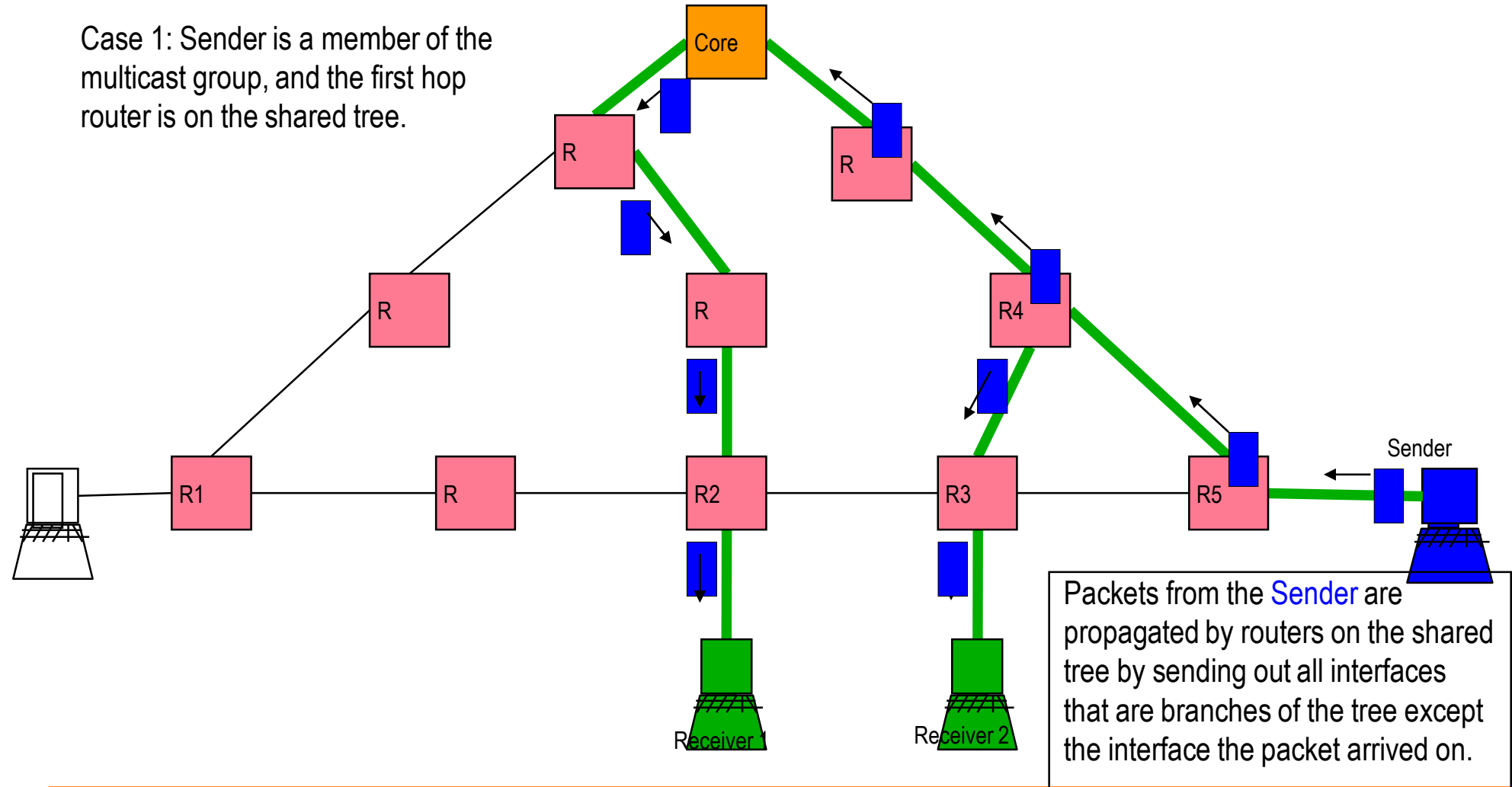
- Unica árvore SPT por grupo => árvores esparsas
  - Adequado a um número grande de emissores
  - As tabelas de encaminhamento escalam bem:
    - dimensão =  $O(\text{Grupos})$
  - Árvores bi-direccionais
-



# Gestão de grupos em CBT



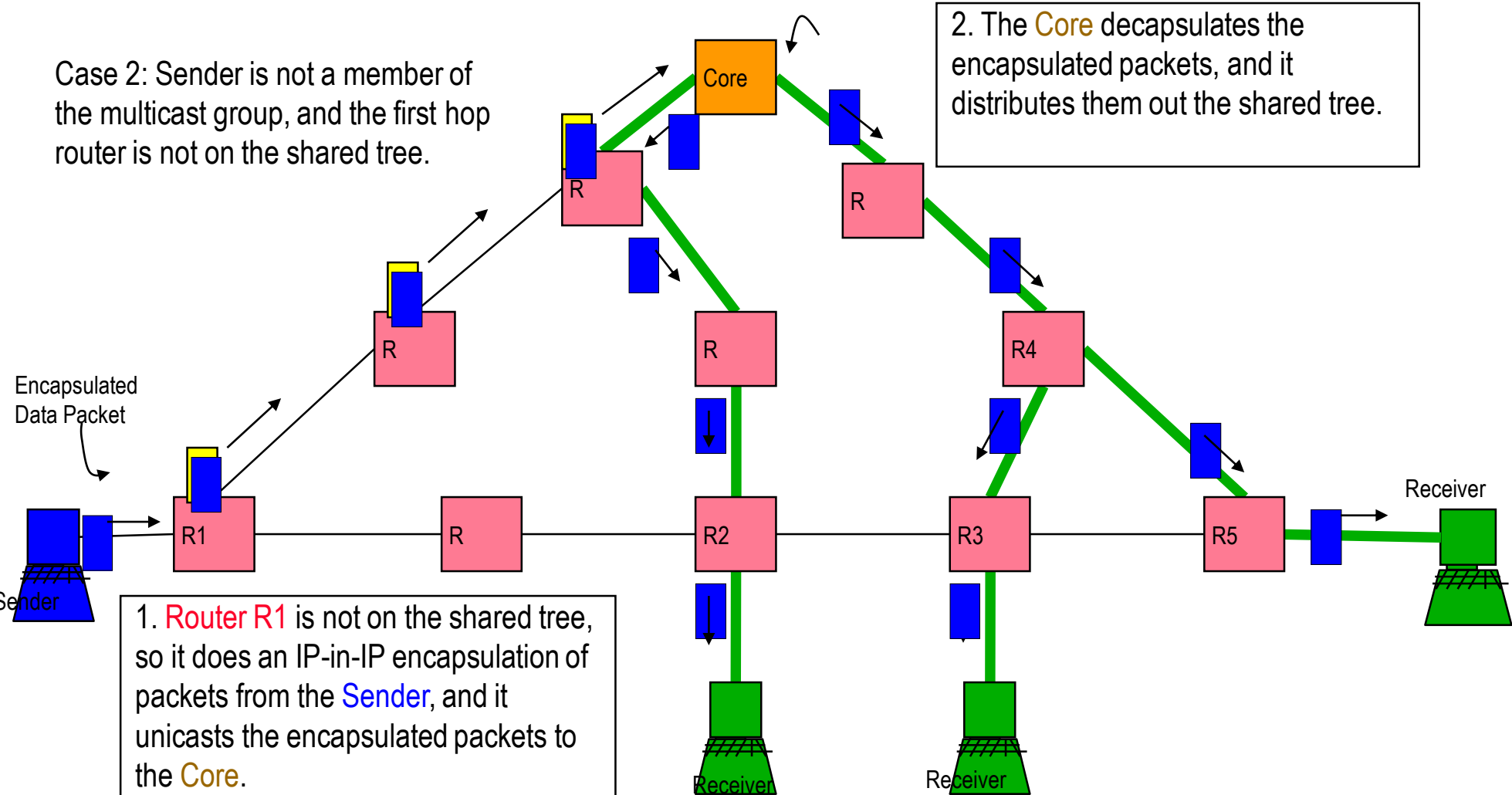
Case 1: Sender is a member of the multicast group, and the first hop router is on the shared tree.



# Envio de dados em CBT (2)



Case 2: Sender is not a member of the multicast group, and the first hop router is not on the shared tree.





- Introdução
- Grupos Multicast: IGMP
- Encaminhamento Multicast
  - Distance Vector Multicast Routing Protocol
  - Core Base Trees
  - Protocol Independent Multicast
  - Border Gateway Management Protocol

# Protocolo Independent Multicast (PIM)

---

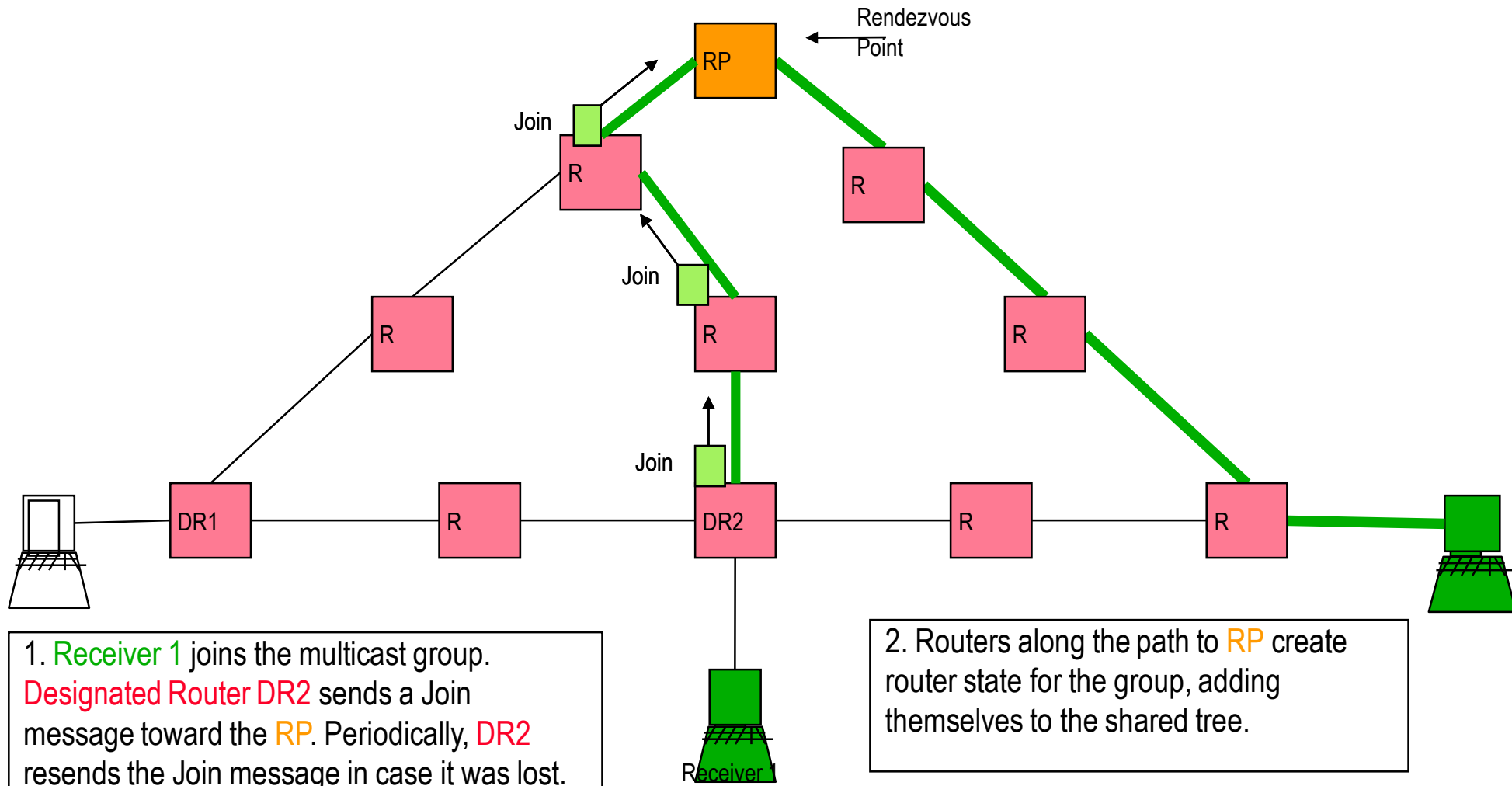


- Usa tabela de encaminhamento *unicast* para a topologia
- Dense mode (PIM-DM)
  - Para grupos com muitos receptores numa região local/global
  - Semelhante ao DVMRP, um algoritmo do tipo *flood and prune*
- Sparse mode (PIM-SM)
  - Para grupos com um pequeno conjunto de receptores amplamente distribuídos
  - Constrói uma **árvore partilhada** por grupo, mas pode construir **árvore centrada** na origem por motivos de eficiência
  - Existe um “join” explícito

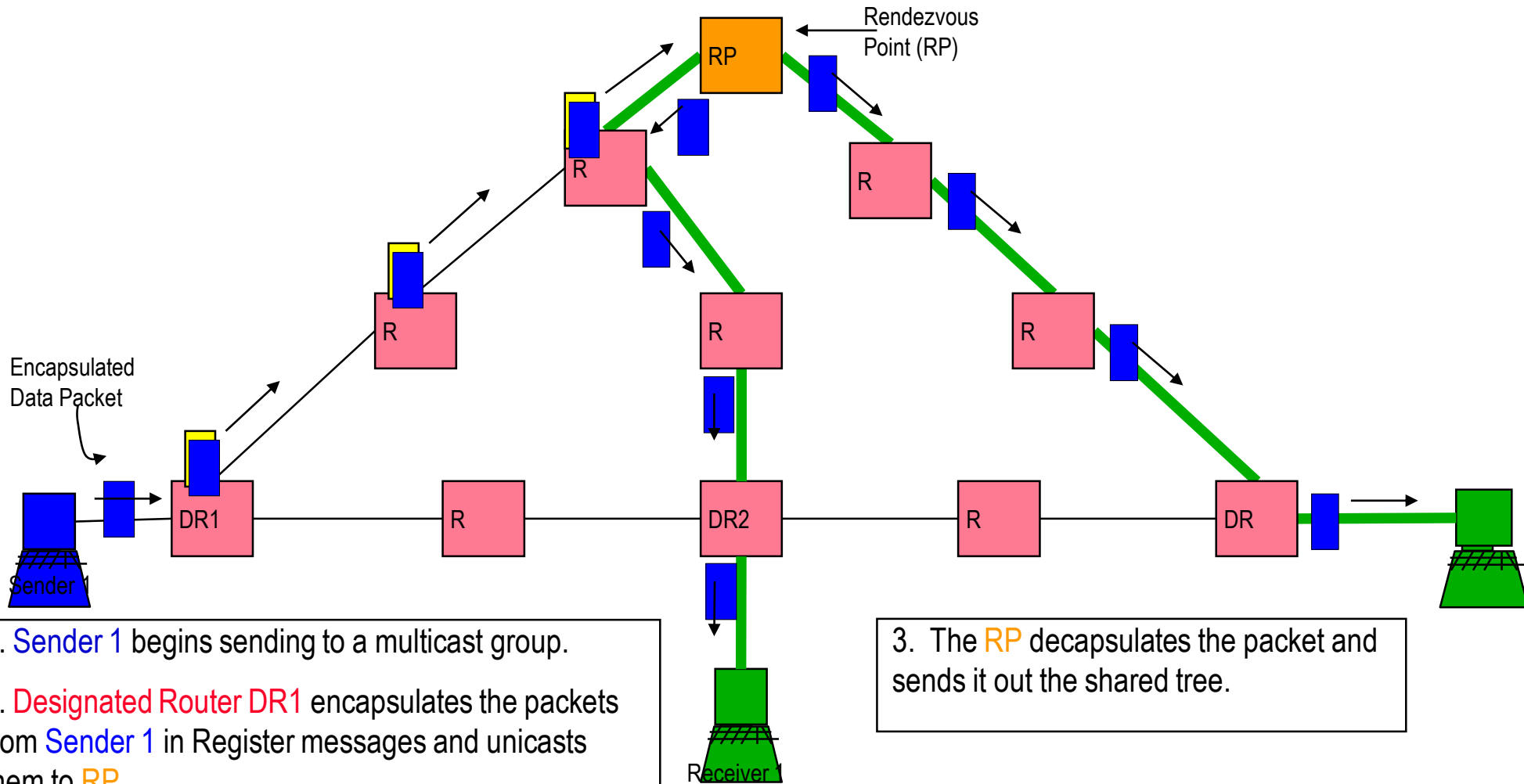


- Protocolo híbrido que combina as características do DVMRP e do CBT
- Adequado a redes amplamente distribuídas e heterógeneas
- A árvore partilhada está centrada no Rendezvous Point (RP)
- A árvore partilhada introduz os emissores aos receptores
- Existem árvores específicas para fluxos de tráfego mais pesados
- Árvore de distribuição unidireccional

# Gestão de grupos em PIM-SM

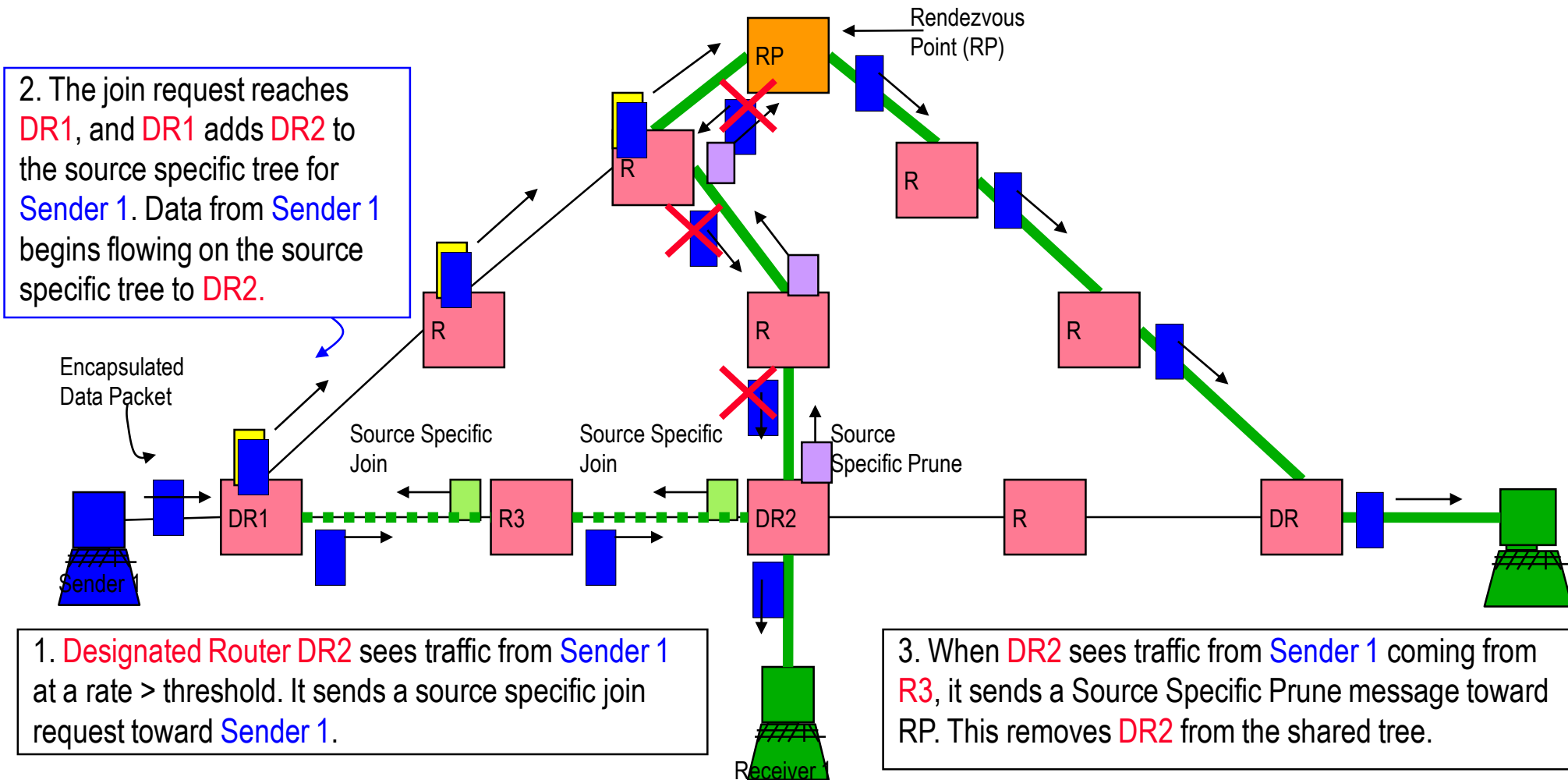


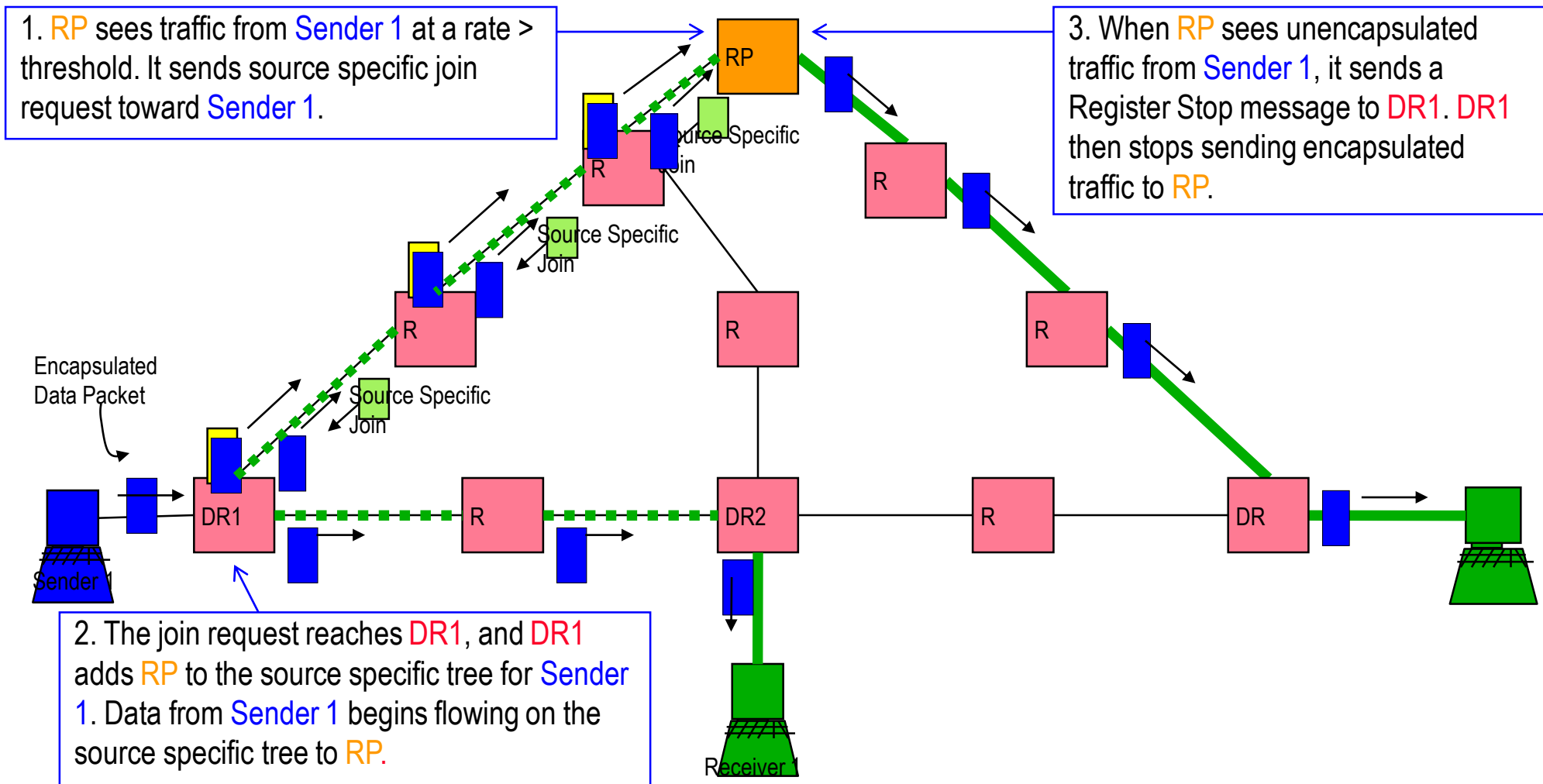
# Envio de dados em PIM-SM





# PIM-SM Específico para uma origem





# PIM Rendezvous Point (RP)

---



- Requisito:
    - Diferentes grupos são mapeados para diferentes RPs
  - Bootstrap Router (BSR)
    - Dinamicamente eleito
    - Constrói um conjunto de endereços IP de RP baseando-se nas mensagens recebidas dos Candidatos-a-RP
  - De que forma os routers conhecem um RP de um dado grupo ?
    - Bootstrap Router broadcasts Bootstrap message com o RP atribuído ao PIM
    - Função de hash no endereço de grupo é mapeada para um RP
-

# Problemas do PIM

---



- Broadcasts globais para todos os Rendezvous Points
  - Sensível à localização do RP
  - Não há controlo administrativo sobre tráfego multicast
  - Concebido como inter domínio, mas agora considerado intra domínio
-



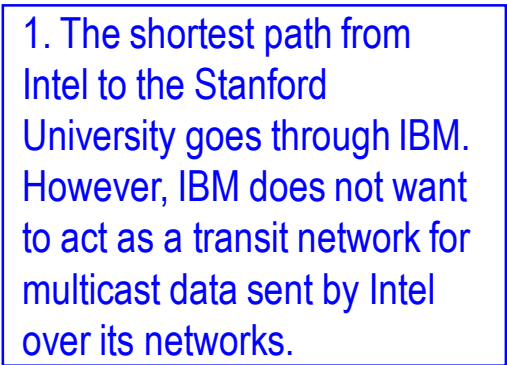
- Introdução
- Grupos Multicast: IGMP
- Encaminhamento Multicast
  - Distance Vector Multicast Routing Protocol
  - Core Base Trees
  - Protocol Independent Multicast
  - Border Gateway Management Protocol

# Border Gateway Multicast Protocol (BGMP)

---



- Motivação
  - Necessidade de hierarquia para encaminhamento multicast
  - Combinar a reserva de endereços multicast e o encaminhamento multicast
  - Protocolos de encaminhamento Inter-AS necessitam de controlo administrativo do tráfego multicast
- Considerações de escalabilidade
  - Necessidade para minimizar o estado no encaminhamento
  - Necessidade de minimizar as mensagens de controlo
  - Enviar dados apenas para onde é necessário

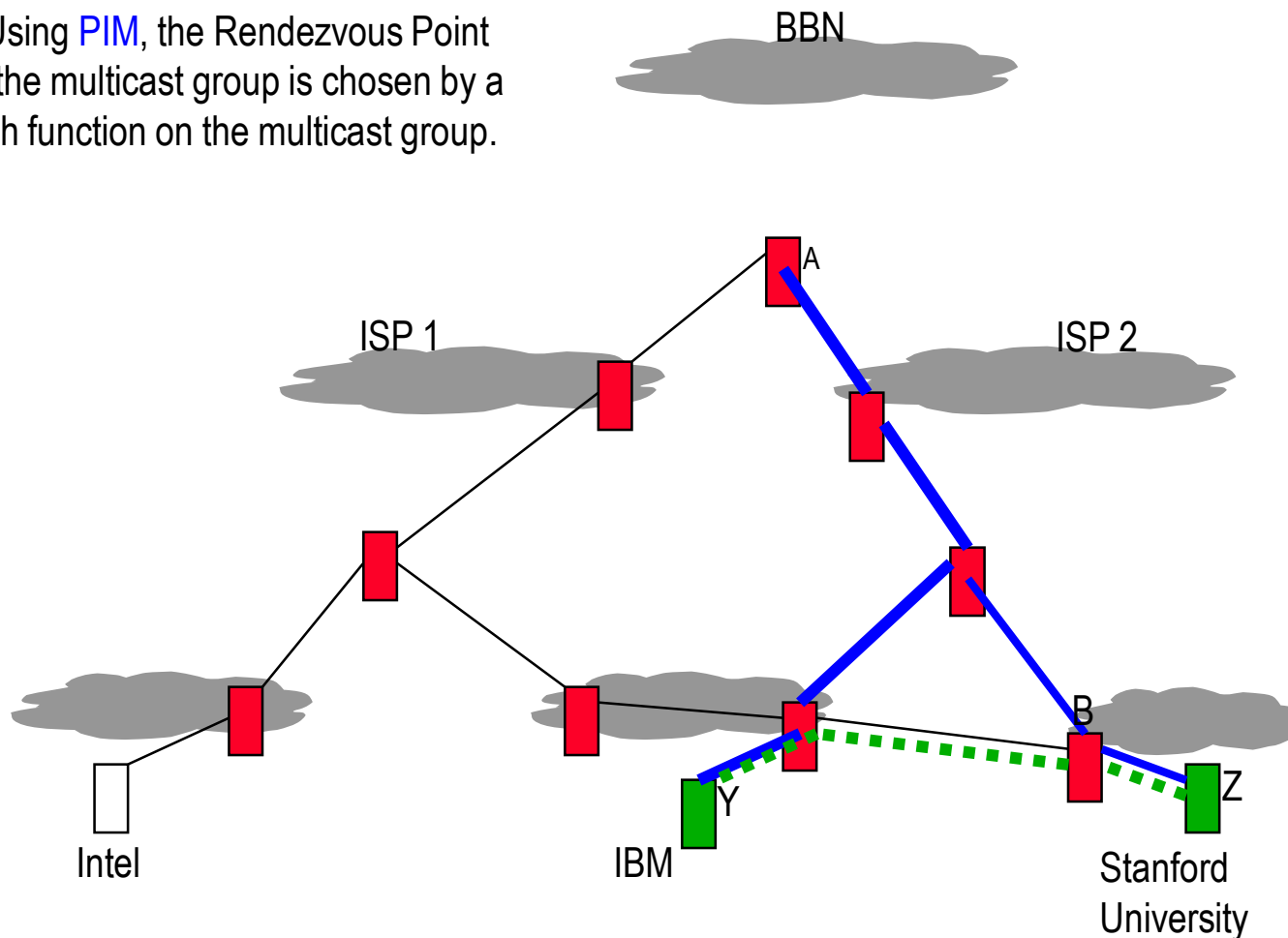


2. IBM installs an administrative policy that does not propagate any multicast routes of Intel senders in outside of IBM's internal network.

# Escolhendo uma árvore partilhada



1. Using **PIM**, the Rendezvous Point for the multicast group is chosen by a hash function on the multicast group.



2. Therefore, the Rendezvous Point for a session started by **Host Z** at the Stanford University might be in BBN at **Router A**. The **PIM shared tree** would cross ISP 2 even though there are no receivers in that direction.

3. If **Host Z** at the Stanford University initiates a conference, the root of the shared tree should be in the Stanford University domain (e.g. **Router B**). The **shared tree** only develops in places with interested receivers downstream.



# Referências

---



- DVMRP: RFC1035
  - CBTP: RFC2189
  - PIM-SM: RFC4601
  - PIM-DM: RFC3973
  - PIM-SSM: RFC3569
  - BGMP: RFC3913
-